

2013

Techniques for Ocular Biometric Recognition Under Non-ideal Conditions

Raghavender Reddy Jillela
West Virginia University

Follow this and additional works at: <https://researchrepository.wvu.edu/etd>

Recommended Citation

Jillela, Raghavender Reddy, "Techniques for Ocular Biometric Recognition Under Non-ideal Conditions" (2013). *Graduate Theses, Dissertations, and Problem Reports*. 652.
<https://researchrepository.wvu.edu/etd/652>

This Dissertation is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Dissertation in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Dissertation has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.

Techniques for Ocular Biometric Recognition Under Non-ideal Conditions

by

Raghavender Reddy Jillela

Dissertation submitted
to the College of Engineering and Mineral Resources
at West Virginia University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Electrical Engineering

Dr. Donald Adjeroh, West Virginia University
Dr. Vijayakumar Bhagavatula, Carnegie Mellon University
Dr. Lawrence Hornak, West Virginia University
Dr. Xin Li, West Virginia University
Dr. Arun Ross, West Virginia University, Chair

Lane Department of Computer Science and Electrical Engineering

Morgantown, West Virginia
2013

Keywords: Biometrics, Ocular, Iris, Face

© 2013 Raghavender Reddy Jillela

Abstract

Techniques for Ocular Biometric Recognition Under Non-ideal Conditions

Raghavender Reddy Jillela

Electrical Engineering, West Virginia University

Dr. Arun Ross, West Virginia University, Chair

The use of the ocular region as a biometric cue has gained considerable traction due to recent advances in automated iris recognition. However, a multitude of factors can negatively impact ocular recognition performance under unconstrained conditions (e.g., non-uniform illumination, occlusions, motion blur, image resolution, etc.). This dissertation develops techniques to perform iris and ocular recognition under challenging conditions. The first contribution is an image-level fusion scheme to improve iris recognition performance in low-resolution videos. Information fusion is facilitated by the use of Principal Components Transform (PCT), thereby requiring modest computational efforts. The proposed approach provides improved recognition accuracy when low-resolution iris images are compared against high-resolution iris images. The second contribution is a study demonstrating the effectiveness of the ocular region in improving face recognition under plastic surgery. A score-level fusion approach that combines information from the face and ocular regions is proposed. The proposed approach, unlike other previous methods in this application, is not learning-based, and has modest computational requirements while resulting in better recognition performance. The third contribution is a study on matching ocular regions extracted from RGB *face* images against that of near-infrared *iris* images. Face and iris images are typically acquired using sensors operating in visible and near-infrared wavelengths of light, respectively. To this end, a sparse representation approach which generates a joint dictionary from corresponding pairs of face and iris images is designed. The proposed joint dictionary approach is observed to outperform classical ocular recognition techniques. In summary, the techniques presented in this dissertation can be used to improve iris and ocular recognition in practical, unconstrained environments.

To My Family

Acknowledgments

This thesis would not have been possible without the support and guidance of many people.

First, I would like to express my deepest gratitude to my advisor, Dr. Arun Ross. I vividly remember my initial interactions with him during the first semester of my Masters, when he taught the *Introduction to Biometrics* course. Little did I know that he would greatly inspire me in pursuing both, Masters and PhD in the field of biometrics. His dedication, patience, and hardworking nature have constantly inspired me over the years. Working with him has taught me to conduct research with rigor, and present my work with clarity. I truly consider it as a privilege to have him as an advisor and instructor throughout my graduate studies.

I would like to thank my committee members for their valuable guidance and suggestions in my research. Dr. Xin Li's courses in image- and video-processing have been extremely helpful in getting started with my research. My thesis has been greatly benefitted by Dr. Donald Adjero's detailed feedback. It was Dr. Lawrence Hornak who inspired me to think practically in research. Special thanks to Dr. Vijayakumar Bhagavatula, who agreed to be a member of my committee on a short notice. His discussions during IARPA meetings and subsequent thesis presentations reminded me to maintain novelty in my work.

I feel lucky to have interacted with a wide number of students and visitors at the Integrated Pattern Recognition and Biometrics (i-PRoBe) lab. I would like to thank Ajita Rattani, Antitza Dantcheva, Asem Othman, Ayman Abaza, Brian DeCann, Cunjian Chen, David Casleton, Emanuela Marasco, Gizem Erdogan, Manisha Sam Sunder, Matthew Monaco, Matthias Vanoni, Nicholas Rymer, Nikhil Burri, Olaoluwa Laseinde, Raghunandan Pasula, Ravindra Gadde, Simona Crihalmeanu, Susan El-Naggar, and Yao-hui Ding for their camaraderie in and out of work.

Thanks to all my friends who supported and encouraged me with their best wishes.

I will forever be indebted to my family for their care, love, and patience. Finally, I thank my wife, Divya for her unconditional love and support during the toughest stages of my student life.

Contents

Acknowledgments	iv
List of Figures	viii
List of Tables	xiv
1 Introduction	1
1.1 Ocular Biometrics	1
1.2 Iris Recognition	3
1.2.1 Anatomy of the Human Iris	3
1.2.2 Significance of the Iris Texture	4
1.2.3 Iris as a Biometric	5
1.2.4 Steps involved in Iris Recognition	7
1.3 Existing Iris Recognition Research	9
1.3.1 Image Acquisition	9
1.3.2 Imaging Wavelengths	11
1.3.3 Iris Image Quality	12
1.3.4 Iris Segmentation	12
1.3.5 Iris Encoding and Matching	14
1.3.6 Iris Recognition under Near-Ideal Conditions	15
1.4 Ocular Recognition	15
1.4.1 Existing Ocular Recognition Research	18
1.5 Factors that affect Iris and Ocular Recognition	19
1.6 Motivation	26
1.7 Contributions	27
1.8 Thesis Organization	29
2 Low Resolution Iris Recognition	30
2.1 Introduction	30
2.2 Image-level Fusion	32
2.2.1 Existing Work	32
2.2.2 Fusion in Iris Videos	32
2.3 Principal Components Transform	34
2.3.1 PCT versus PCA	34
2.3.2 Mathematical Formulation	35
2.4 Proposed Approach	38
2.5 Experiments and Results	39
2.5.1 Database	39

2.5.2	Pre-processing	40
2.5.3	Iris Segmentation	40
2.5.4	Recognition Accuracy	42
2.5.5	Proposed Approach on High-Resolution Images	45
2.5.6	Effect of Down-sampling Methods	47
2.6	Score-Level Fusion	49
2.7	Effect of Image Registration	50
2.8	Summary	51
3	Plastic Surgeries: Face & Ocular Biometric Fusion	53
3.1	Introduction	53
3.2	Facial Plastic Surgeries	54
3.3	Existing Approaches	56
3.4	Motivation	56
3.5	Ocular Recognition	58
3.6	Proposed Approach	59
3.6.1	Scale Invariant Feature Transform	60
3.6.2	Local Binary Patterns	61
3.6.3	Score-level Fusion	61
3.7	Database	62
3.7.1	Face Dataset A	63
3.7.2	Face Dataset B	64
3.7.3	Ocular Dataset	64
3.8	Experiments and Results	65
3.8.1	Face recognition performance	65
3.8.2	Ocular recognition performance	65
3.8.3	Score-level fusion performance	65
3.8.4	Effect of individual surgeries	67
3.9	Summary	70
4	Ocular Matching in Face and Iris Images	71
4.1	Introduction	71
4.1.1	Motivation	72
4.2	Problem Significance	74
4.3	Database	75
4.3.1	Challenges	77
4.4	Outline of Experiments	79
4.5	Iris Recognition	79
4.5.1	Open Source Algorithm	80
4.5.2	Commercial Algorithm	83
4.6	Ocular Recognition	84
4.6.1	Baseline - Local Binary Patterns	85
4.6.2	Normalized Gradient Correlation	85
4.7	Ocular Matching using Joint Dictionary Approach	89
4.7.1	Sparse Representation Framework	89
4.7.2	Joint Dictionary Approach	90
4.7.3	Dictionary Learning and Matching	91
4.8	Computational Details	94

<i>CONTENTS</i>	vii
4.9 Score-level fusion	94
4.10 Summary	95
5 Summary and Conclusions	97
5.1 Research Contributions	97
5.2 Directions for Future Research	98
APPENDICES	100
A Methods for Iris Segmentation	101
A.1 Classical Iris Segmentation Algorithms	101
A.1.1 Integro-Differential Operator	101
A.1.2 Hough Transform	103
A.2 Other Prominent Approaches	105
A.2.1 Geodesic Active Contours	106
A.2.2 Variational Level Sets	110
A.2.3 Fourier-based Approximation	114
A.3 Role of Image Acquisition on Iris Segmentation	117
A.4 Segmenting Visible Wavelength Iris Images	123
A.5 Performance Evaluation of a Few Iris Segmentation Techniques	126
A.6 Approaches to Refine Iris Segmentation	127
A.7 Predicting Errors in Iris Segmentation	131
B Database Release Agreement	135
Bibliography	137

List of Figures

1.1	Ocular biometric traits: (a) retinal vasculature pattern, (b) iris texture, (c) conjunctival vasculature pattern, (d) ocular region, (e) oculomotor characteristics, and (f) complex eye movements. Image in panel (e) taken from [1].	3
1.2	Schematics showing (a) the anatomy of the eye, and (b) the external structure of the iris. Image in (a) taken from [2].	5
1.3	Images showing (a) the non-circularity, and (b) nasal inclination of the iris. Notice that in (b), the pupil is slightly inclined towards the subjects nose (on the right side of the image). The distance from the pupillary boundary to the limbus boundary on the left side of the image, d_1 , is greater than the distance on the right side, d_2	6
1.4	Close-up view of an iris, showing its complex texture. Image taken from [3].	6
1.5	A block diagram of an iris recognition system.	9
1.6	Examples of the three different types of iris sensors: (a) Retica Mobile-Eyes, (b) Oki IrisPass and (c) Iris on the Move (IOM) Passport system.	10
1.7	Output obtained by applying an integro-differential operator to detect both the iris and eyebrow boundaries.	13
1.8	Anatomical characteristics and features that contribute to the uniqueness of ocular region.	16
1.9	Sample situations where ocular information can be used to improve (a) face, and (b) iris recognition performance, respectively.	17
1.10	Ocular information could be acquired with minimal privacy violation. . . .	18
1.11	Sampling point patterns for (a) local (e.g., SIFT) and (b) global (e.g., GOH) ocular feature extraction schemes.	19
1.12	(a) Iris, and (b) ocular image acquired under poor illumination conditions. Although the iris boundaries or location can be determined by humans, automatic recognition using such images is extremely difficult [4].	20
1.13	An iris image containing specular reflections on the pupillary boundary. . .	20
1.14	Closeup of an iris image acquired at a large stand-off distance.	20
1.15	Sensor noise in an (a) iris and (b) ocular image.	21
1.16	An iris image showing occlusions caused by the eyelids.	22
1.17	An iris image showing occlusions due to eyelashes.	22
1.18	Iris image of a subject suffering from congenital iris abnormality. Notice that the pupil boundary is neither circular nor elliptical. Image taken from [5].	23
1.19	(a) An iris image of a subject wearing contact lens. (b) Ocular image of a subject wearing eye glasses. Image in (a) taken from [6] ©Elsevier.	23
1.20	An (a) iris and (b) ocular image containing motion blur.	24
1.21	An off-angled iris image.	24

1.22	Images showing the variation in appearance of ocular region: (a) without, and (b) with deformations, respectively. Image source: 04233d1632-08-l.jpg and 04233d1649-03-l.jpg from [7].	25
1.23	An iris image with artifacts caused by severe image compression. Image taken from [8].	25
1.24	Sample scenarios that motivate the need for rigorous research in ocular recognition: (a) recognition at a distance (image from [9]), (b) covert surveillance (image from [10]), (c) cross-sensor matching (image from [11]), (d) cross-source matching (author's personal images), and (e) spoofing (image from [12]).	28
2.1	Panels (a), (b), and (c), respectively, show the face, the eye, and the normalized iris regions (in the visible spectrum) of a subject standing close to the sensor. The corresponding regions for a subject with larger stand-off distance are shown in panels (d), (e), and (f), respectively.	31
2.2	Difference between PCA and PCT: (a) In an MN dimensional space, the conventional PCA approach considers each image as a point (denoted by a circle). (b) In an n dimensional space, the PCT approach considers every pixel intensity vector as a point (denoted by a diamond).	35
2.3	The process of stacking image pixels used to generate image data matrix \mathbf{X} for the PCT approach.	36
2.4	Left column: Normalized input iris images. Middle column: principal components of the input data arranged in the decreasing order of magnitude. Right column: reconstructed output obtained by using the PCT approach with $q = 2$. Images have been scaled to fit the document.	37
2.5	Proposed image-level fusion scheme.	39
2.6	A sample right iris frame of resolution 640×480 down-sampled to resolutions of 320×240 , 160×120 , and 80×60 using the image averaging operator. Images have been scaled to fit the document.	40
2.7	Normalized probe images corresponding to an iris frame at multiple resolutions: (a) 640×480 , (b) 320×240 , (c) 160×120 , and (d) 80×60 . Images have been scaled to fit the document.	42
2.8	ROC curves for the right iris data, obtained by matching probe sets with the gallery.	43
2.9	ROC curves for the right iris data, obtained before and after applying the proposed technique. Note that by applying the proposed approach, the Equal Error Rate of <i>ProbeSet L4</i> is reduced from 6.09% to 1.76%.	44
2.10	Genuine and impostor match score distributions for the right iris data: (a) before, and (b) after applying the proposed technique.	45
2.11	ROC curves for right iris <i>ProbeSet L1</i> and <i>ProbeSet L4</i> , before and after applying the proposed approach.	46
2.12	A sample right iris frame of resolution 640×480 down-sampled to resolutions of 320×240 , 160×120 , and 80×60 using the Gaussian filtering approach. Images have been scaled to fit the document. Note that the same image was used in Figure 2.6.	48
2.13	Normalized images obtained from <i>ProbeSet L1</i> to <i>L4</i> that are downsampled using Gaussian ((a), (c), (e), and (g)), and averaging ((b), (d), (f), and (h)) operators. Images have been scaled to fit the document.	49

2.14	Normalized images corresponding to a single subject from <i>ProbeSet L4</i> before (left column) and after (right column) applying a random 10 pixel shift. Images have been scaled to fit the document.	51
3.1	Some of the major facial plastic surgeries. Image taken from the FRGC database [13].	55
3.2	Images showing the degree to which the appearance of a human face can be modified by plastic surgeries. Top row: (a) before and (b) after a minor plastic surgery (blepharoplasty). Bottom row: (c) before, and (d) after multiple plastic surgeries.	55
3.3	Facial images of a subject (a) before, and (b) after undergoing rhytidectomy. (c) and (d): Corresponding ocular images of the same subject. Note that the variation in the appearance of the face, from a visual perspective, is much larger than that of the ocular region.	58
3.4	Sample bi-ocular images used in this work. Note that the images have been resized for the purpose of clarity.	59
3.5	Neighborhood for computing the LBP of pixel p	61
3.6	A schematic representation of the proposed approach.	62
3.7	Images exhibiting some of the challenges in the facial plastic surgery database. (a) and (d): images with varying resolution, scale and inter-ocular distances corresponding to the same subject. (b) and (e): variations in expressions of a subject. (c) and (f): duplicate entries. The image in (c) is listed as ID #26300 and its duplicate image in (f) is re-listed as ID #28519. Note the difference in identification labels, although they belong to the same subject who has undergone multiple surgeries. This incorrect labeling can negatively impact the perceived matching accuracy.	63
3.8	CMC curves showing the recognition performances of VeriLook and PittPatt on <i>face dataset A</i> and <i>face dataset B</i>	66
3.9	CMC curves showing the recognition performances of LBP and SIFT on <i>ocular dataset</i>	67
3.10	CMC curve showing the recognition accuracies obtained using score-level fusion of face scores, ocular scores, and a combination of the two.	69
4.1	Sample images corresponding to (a) face, and (b) iris modalities, respectively.	72
4.2	Closeup of the left-side ocular regions from (a) face, and (b) iris images shown in Figure 4.1. The goal is to perform ocular matching using the regions observed in (a) and (b). Notice the variations in scale, resolution, image acquisition wavelength, viewing angle, and the level of detail. The average number of pixels across the irides in (a) and (b) are 35 and 110 pixels, respectively.	72
4.3	An illustration depicting the different modalities in databases. Note that relating the identities stored in <i>Database C</i> with those stored in other databases can be a challenging process.	73
4.4	A sample scenario depicting the need for ocular matching in face-iris image pairs.	74
4.5	Sample images from the BioCoP database showing the (a) face, (b) cropped ocular region, and (c) iris of a subject. <i>Subject's approval to use the images for illustration purposes is on file</i>	76

4.6	Corresponding ocular regions from (a) a face image and (b) an iris image acquired under visible and NIR spectra, respectively. Notice the variation in textural appearance of the iris within the images.	77
4.7	Images showing variations in the viewing angle between corresponding ocular regions obtained from (a) a face image and (b) an iris image, respectively.	78
4.8	Variations in illumination observed in images acquired by the NIR iris sensor.	78
4.9	Occlusions of the iris and ocular regions caused by (a) eyelids, (b) eyelashes and (c) hair, as observed in the images acquired by the iris sensor.	78
4.10	Images depicting the sensor noise and non-uniformity in imaging.	79
4.11	Sample NIR iris (top row) and VIS ocular region (bottom row) images showing correct iris segmentation output obtained using: Integro-Differential Operator [(a) and (d)], Hough transform [(b) and (e)], and Geodesic Active Contours [(c) and (f)] based algorithms.	81
4.12	Sample NIR iris (top row) and VIS ocular region (bottom row) images showing incorrect iris segmentation output obtained using: Integro-Differential Operator [(a) and (d)], Hough transform [(b) and (e)], and Geodesic Active Contours [(c) and (f)] based algorithms, respectively.. . . .	81
4.13	(a) Sample <i>NIR</i> iris image. (b) Corresponding VIS ocular region cropped from an <i>RGB</i> face image. (c), (d), and (e) are R, G, and B channel images extracted from the VIS ocular image, respectively.	82
4.14	ROC curves obtained by matching iris regions extracted from R, G, and B channels of the VIS ocular images with those extracted from the NIR iris images, using Libor Masek's open source implementation. <i>Note that these curves correspond to matching performance obtained using a subset of images (100 NIR iris and 100 VIS ocular images).</i>	83
4.15	ROC curves obtained using VeriEye to match irides from left-side NIR iris images and left-side VIS ocular images.	84
4.16	ROC curves corresponding to the ocular region matching using LBP on left-side images.	87
4.17	ROC curves corresponding to ocular matching using the Normalized Gradient Correlation technique on left-side images.	88
4.18	ROC curves corresponding to ocular matching using the joint dictionary based sparse representation approach on left-side images.	94
4.19	ROC curves for the cross-spectral ocular matching using (a) LBP, (b) NGC, (c) proposed joint dictionary based sparse representation approach, and (d) weighted score-level fusion of all the considered techniques.	96
A.1	Iris segmentation using the integro-differential operator.	102
A.2	Output obtained by applying an integro-differential operator to detect both the iris and eyebrow boundaries.	103
A.3	(a) An eye image in which the limbus boundary is occluded by the eyelashes. (b) Approximate limbus boundary obtained using Daugman's integro-differential operator. (c) Precise limbus boundary obtained by a curve evolution technique. Notice that the precise segmentation helps in avoiding the noise caused by eyelash occlusion.	106
A.4	Stopping function for the geodesic active contours. (a) Original iris image, (b) stopping function K , and (c) modified stopping function K'	108

A.5	Contour initialization for iris segmentation using GAC. (a) Zeroth level set (initial contour), (b) mesh plot denoting the signed distance function ψ . . .	109
A.6	Evolution of the geodesic active contour during iris segmentation. (a) Iris image with initial contour, (b) embedding function ψ (X and Y axes correspond to the spatial extent of the eye image and the Z axis represents different level sets), (c,d,e,f) contours after 600 and 1400 iterations, and their corresponding embedding functions, and (g,h) Final contour after 1800 iterations (contours shown in white).	111
A.7	The final contour obtained when segmenting the iris using the GAC scheme. (a) Example of a geodesic contour splitting at various local minima, (b) final contour (contours shown in white).	112
A.8	Rough contours of the pupillary and limbus boundaries obtained using the elliptical model. Notice that the rough contours do not precisely match the true iris boundaries. Image source: Roy et. al. [14] ©Elsevier	112
A.9	Final output obtained using the variational level set approach, where both the iris boundaries are precisely determined. Image source: Roy et. al. [14] ©Elsevier	115
A.10	Output of iris segmentation scheme using variational level set approach on some non-ideal iris images from the UBIRIS Version 1 dataset. Notice that the segmentation results are fairly precise. Image source: Roy et. al. [14] ©Elsevier	115
A.11	Iris segmentation using Daugman's Fourier approximation approach. Image source: Daugman [15] ©IEEE. Image has been edited for clarity.	117
A.12	Conventional iris image acquisition system requiring considerable user cooperation. Image source: http://www.life.com/image/1668585	118
A.13	Image acquisition setup for the public use iris recognition system. Image source: Negin et. al. [16] ©IEEE	119
A.14	Image of a user's torso acquired by the WFOV camera. Image source: Negin et. al. [16] ©IEEE	119
A.15	Image of the user's eye, acquired by the NFOV camera. Image source: Negin et. al. [16] ©IEEE	120
A.16	Left: Circular grid used for iris localization. Right: Iris image with the circular grid overlaid. Image source: Negin et. al. [16] ©IEEE	120
A.17	Image acquisition setup for the iris on the move system. Image source: Matey et. al. [9] ©IEEE	121
A.18	Image acquired using the IOM system, exhibiting a specular reflection pattern that corresponds to the illuminator pattern. Image source: Matey et. al. [9] ©IEEE	122
A.19	Iris segmentation using the specular reflection pattern in the image. Once the specular reflection pattern is detected, a specified area around the pattern is used for unwrapping and feature extraction. Image source: Matey et. al. [9] ©IEEE	123
A.20	An iris image acquired in the visible wavelength. Notice that the intensity contrast between the pupil and the iris is low.	124
A.21	Left: Over segmented limbus boundary. Right: Under segmented limbus boundary.	128

A.22 Unwrapped iris images containing eyelash occlusion. Image source: Zhang et. al. [17] ©IEEE	128
A.23 Left: x derivative for a Sobel Edge filter. Center: image region under consideration. Right: y derivative for a Sobel edge filter.	129
A.24 Top: Image affected by eyelash occlusion. Bottom: Output obtained by applying the eyelash removal technique. Image source: Zhang et. al. [17] ©IEEE	130
A.25 Splitting of the eyelid into four portions to allow better detection of eyelid. Image source: Liu et. al. [18] ©IEEE	131
A.26 Left: Eyelid detection using Libor Masek's approach. Right: Improved eyelid detection using the proposed modifications. Image source: Liu et. al. [18] ©IEEE	131

List of Tables

2.1	Segmentation accuracies of the three techniques using ProbeSet L4.	41
2.2	Equal Error Rates obtained using the left iris video data.	44
2.3	Equal Error Rates obtained before and after applying the proposed approach on right and left iris <i>ProbeSet L1</i> data.	47
2.4	Equal Error Rates obtained before and after applying the proposed technique on right and left iris <i>ProbeSet L1</i> data downsampled by the Gaussian operator.	48
2.5	Equal Error Rates obtained before and after applying score-level fusion on the right and left iris video data.	50
2.6	Equal Error Rates obtained after image shifting and after applying the proposed technique on right iris <i>ProbeSet L4</i> data.	51
3.1	List of algorithms used for performing face recognition on plastic surgery images and the corresponding rank-one accuracies.	57
3.2	List of major facial plastic surgeries separated by the corresponding regions whose appearance can be potentially affected.	58
3.3	Number of images used in each dataset, along with their spatial resolutions.	64
3.4	Duplicate image pairs that reduce the recognition performance at rank-one. Notice the difference in the identification tags, that causes the genuine pairs to be reckoned as impostors.	68
3.5	Example face and ocular image pairs (pre- and post-surgery) that were not correctly matched at rank-1 by the face recognition systems, but were correctly matched at rank-1 after performing fusion.	68
3.6	Rank-one recognition accuracies corresponding to individual surgeries obtained using the face, ocular, and fusion schemes on images from face dataset B and the ocular dataset.	69
4.1	Publications studying the impact of varying imaging factors on iris and ocular recognition (listed chronologically).	75
4.2	Specifications of the considered BioCoP database [19].	76
4.3	List of major biometrics research publications focused on the ocular region	86
4.4	Equal Error Rates obtained using left-side images of the considered BioCoP database.	93
4.5	Equal Error Rates corresponding to different values of λ	95

A.1 Segmentation accuracies of the three techniques evaluated in this work.	127
---	-----

Chapter 1

Introduction

सर्वेन्द्रियाणम् नयनम् प्रधानम्

sarvēndriyānām nayanam praDhānam

Translated from Sanskrit: Of all the senses, vision
- imparted by the **eyes** - is the most important.

1.1 Ocular Biometrics

The term ocular, derived from its Latin root *oculus*, broadly refers to the anatomical regions and structures related to the eye. *Ocular biometrics* refers to the recognition (identification or verification) of individuals using the information offered by ocular modalities. The public perception of ocular biometrics has been largely limited to iris recognition. However, research suggests that several other ocular modalities could serve as biometric indicators. Following is a list of ocular modalities that have been used as biometric traits:

1. **Retina:** The fractal-like growth of the retinal vascular pattern can be used as a reliable biometric trait [20]. Such unique blood vessel patterns can be captured using an image acquisition system comprising of a retina illuminator and a suitable camera. A sample image obtained using such system is shown in Figure 1.1(a). A major limitation of retinal recognition is that it requires a significant amount of user cooperation during image acquisition to achieve reliable recognition accuracy.
2. **Iris:** The multi-layered nature of the iris provides it with complex textural patterns on its surface. These textural variations are very distinctive [21], and render iris a very reliable biometric in general. A sample iris image showing its complex texture

pattern is provided in Figure 1.1(b). Currently, iris recognition is one of the most active areas of biometric research.

3. **Conjunctival vasculature:** The blood vessel patterns observed in the sclera (white) of the eye have been proven to be a biometric trait under constrained conditions [22]. Similar to retina, imaging the conjunctival vasculature patterns requires considerable user cooperation. Figure 1.1(c) shows a sample conjunctival vasculature pattern observed in an eye.
4. **Ocular region:** Ocular region* refers to a rectangular region of fixed size, immediately surrounding the eye globe (or, eye socket). Depending on the area of the region considered, ocular images typically contain anatomical features such as the iris, conjunctiva (white of the eye), eyelashes and the eyelids (if the eye is closed), eyebrows and moles/scars around the eye (if present). Research indicates that the information extracted from the ocular region can serve as a soft biometric trait [24]. The utility of this trait is especially pronounced when the eye is closed and the iris information cannot be acquired. A sample ocular image is shown in Figure 1.1(d).
5. **Oculomotor Plant Characteristics:** The oculomotor plant consists of the non-visible, internal eye globe, its surrounding tissues, ligaments, muscles and tendon-like components [1]. Research indicates that the dynamics of an oculomotor plant, in response to a controlled visual stimulation, can serve as a biometric trait [25]. However, acquiring such information is feasible only under highly constrained conditions. A diagram of the internal muscles of the eye that contribute to the oculomotor plant characteristics is provided in Figure 1.1(e).
6. **Complex Eye Movements:** The fixation and saccadic movement information of an eye can be quantized by various features, and used for identification purposes [26]. Examples of such features include fixation count, saccade velocity, saccade amplitude, scanpath length, etc. Much like oculomotor plant characteristics, complex eye movement information can be acquired only under highly constrained conditions. Figure 1.1(f) shows a sample visual stimuli overlaid with some of the CEM features.

*In some existing research [23], this region is also referred as the *periocular region*.

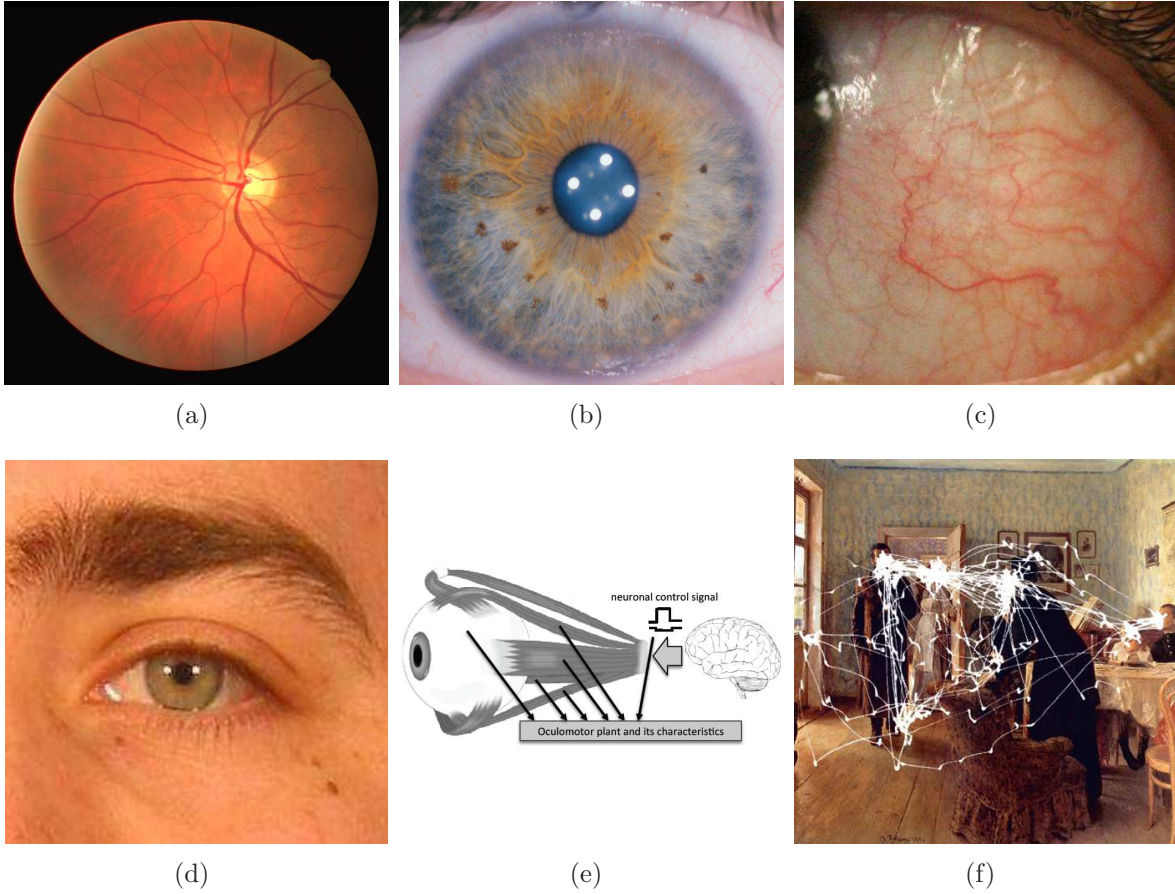


Figure 1.1: Ocular biometric traits: (a) retinal vasculature pattern, (b) iris texture, (c) conjunctival vasculature pattern, (d) ocular region, (e) oculomotor characteristics, and (f) complex eye movements. Image in panel (e) taken from [1].

It has to be noted that the first four modalities in the above list (i.e. retina, iris, conjunctival vasculature and ocular) are considered as *anatomical* traits. In contrast, the latter two are considered to be combination of *physiological* and *behavioral* traits [25].

1.2 Iris Recognition

Of all the ocular biometric traits listed in Section 1.1, iris is considered to be the most reliable. This is based on its uniqueness, performance and circumvention properties. In the following subsections, the fundamentals of iris recognition are discussed in detail.

1.2.1 Anatomy of the Human Iris

The word iris (pl. irides) is derived from the Greek word for mythological goddess of rainbow, *iris* [27]. The word was used for any brightly colored circle [28], and therefore,

was also used to describe the colored portion of the exterior eye. From an anatomical perspective, iris refers to the annular region within the eye, that is located just behind the cornea and in front of the lens [29]. The structure of the iris is that of a thin diaphragm, in the shape of a 3D truncated cone. Such structure is imparted by its alignment with the lens of the eye [30]. A diagram showing the location of the iris in an eye is provided in Figure 1.2(a).

The anterior surface of the iris is divided into two regions: the pupillary zone, and the ciliary zone. The pupillary zone is the inner region, that is located closer to the pupil. The ciliary zone is the outer region, that comprises the rest of the iris. These two regions are separated by the collarette. The iris is at its thickest at the collarette, and thins away radially from the pupil. The average radius of the human iris is approximately 6mm, with an average thickness of about 0.5mm. The functionality of the iris is to regulate the amount of light that enters the eye, by controlling the size of the pupil. The size of the pupil is controlled using the dilator and sphincter muscles, that are connected by stroma. The stroma is a pigmented fibrovascular tissue. A sample image showing the external structure of the iris is provided in Figure 1.2(b).

In most of the iris recognition literature, iris boundaries are considered to be perfectly circular, or elliptic. This assumption helps in easily discarding non-iris regions, when performing automatic iris recognition. However, the iris is not always perfectly circular, or elliptic in nature. An image showing a non-circular iris is shown in Figure 1.3(a). Furthermore, it has to be noted that the pupil exhibits a slight nasal inclination. Therefore, it is not necessary that the center of the pupillary boundary always coincides with that of the limbic boundary. An image showing the nasal inclination of the iris is shown in Figure 1.3(b).

1.2.2 Significance of the Iris Texture

The iris begins to form during the third month of the gestation, and develops a distinctive structure by the eighth month [31]. The distinctiveness is caused by the presence of fibrous and cellular structures such as ligaments, furrows, crypts, rings, frills, corona, collarette, and sometimes moles, freckles, nevi, and other macro-features [32]. The overall appearance of the iris can be roughly described as a multilayered, tangled mesh-like struc-

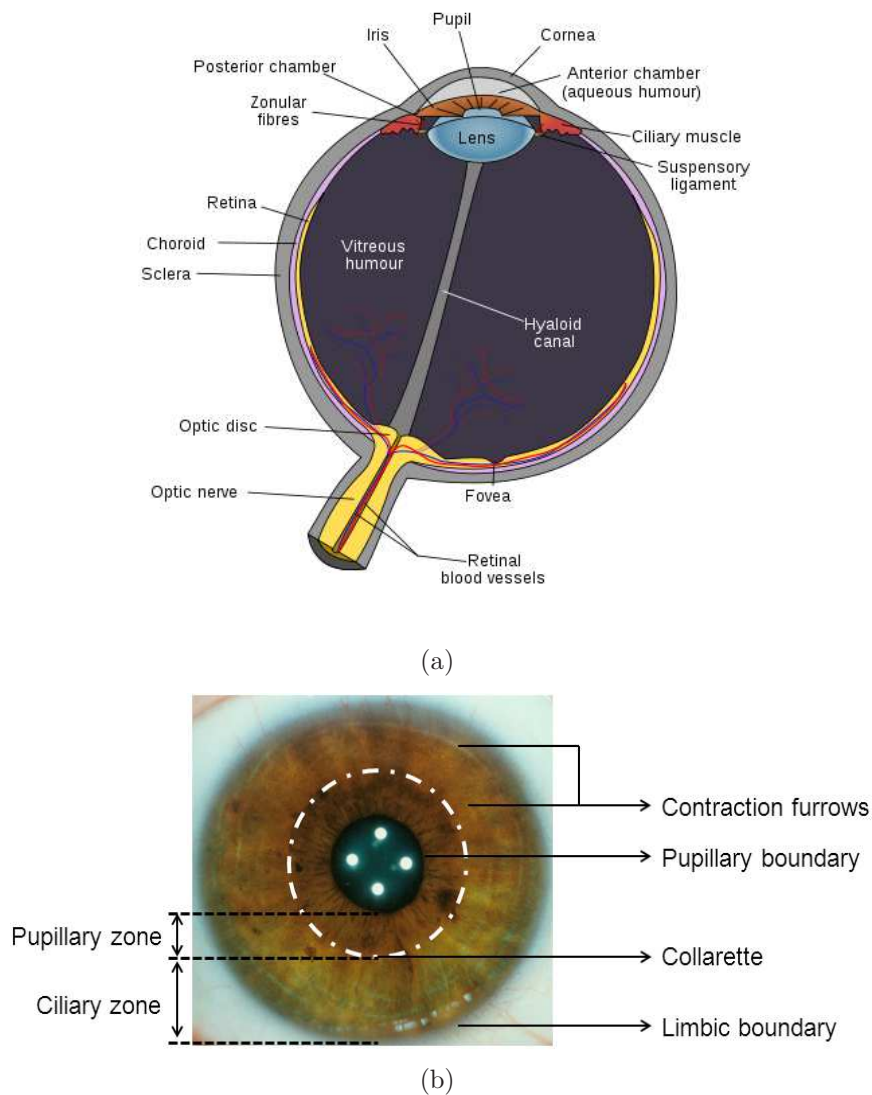


Figure 1.2: Schematics showing (a) the anatomy of the eye, and (b) the external structure of the iris. Image in (a) taken from [2].

ture, which imparts a highly complex texture to its surface. Figure 1.4 provides a close-up view of texture of a sample iris. It is such highly complicated texture that makes the iris a unique biometric.

1.2.3 Iris as a Biometric

The assumed uniqueness of the iris has paved for its usage as a biometric trait for human recognition. A majority of the research in iris recognition literature cites the research by Daugman [33] as a fundamental work in establishing iris as a biometric. However, numerous prior research efforts exist, that allude to the usability of iris for biometric purposes. The

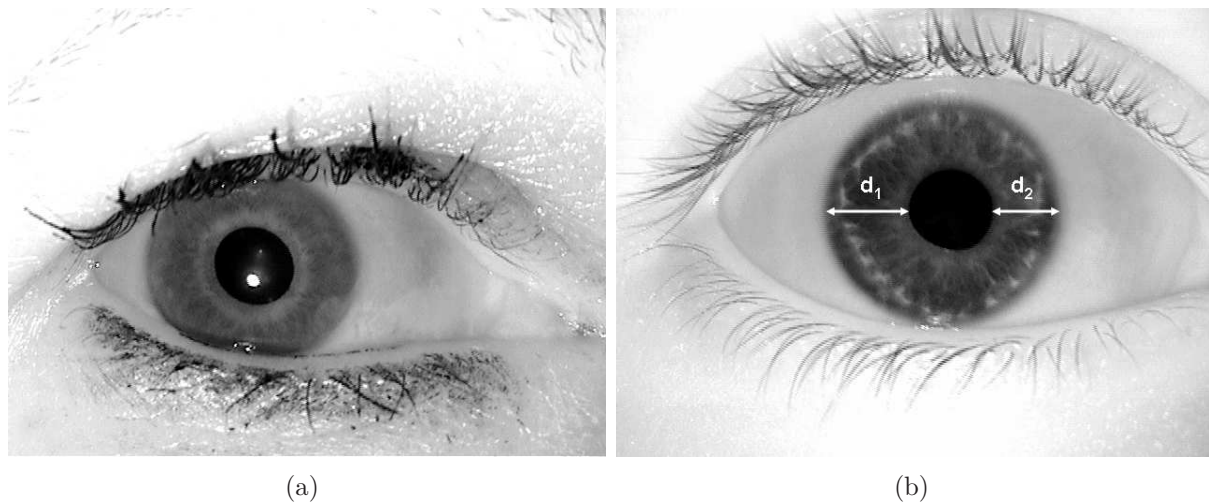


Figure 1.3: Images showing (a) the non-circularity, and (b) nasal inclination of the iris. Notice that in (b), the pupil is slightly inclined towards the subjects nose (on the right side of the image). The distance from the pupillary boundary to the limbus boundary on the left side of the image, d_1 , is greater than the distance on the right side, d_2 .

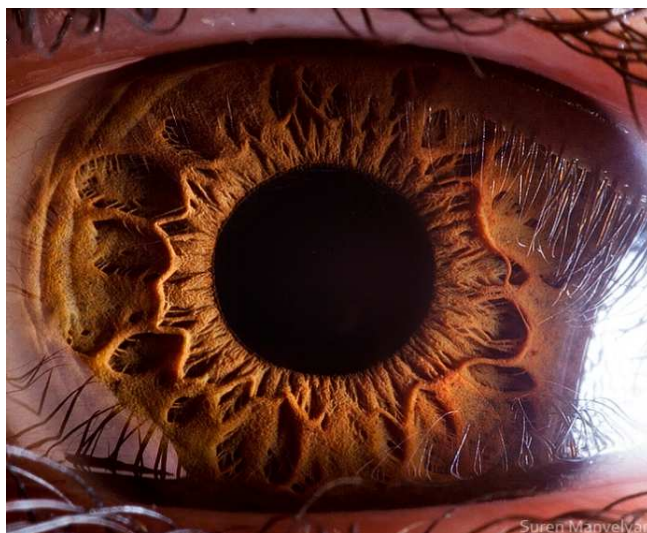


Figure 1.4: Close-up view of an iris, showing its complex texture. Image taken from [3].

first recorded usage of information related to the eye for identification purposes is associated with Alphonse Bertillon. In his work [34, 35], Bertillon describes the usage of eye color patterns to distinguish criminals. Such analysis was based on the notable difference between the eye colors, and their subtle sub-divisions. Davenport and Davenport [36] discussed the ethnographic diversity of the iris, along with its heritability property. Mann's work [37] suggests that the iris is a unique anatomical entity, and its general structure is determined genetically. Rohen and Unger [38] studied the mesh-like structure of the iris, and discussed the stability of iris texture over time. Adler's work [29] in ophthalmology describes that

the texture of the iris is highly detailed and unique. Based on his observations from clinical photographs of irides spanning decades of time span, Adler also suggested that the texture of the iris is very stable.

Wasserman’s research [39] suggests that the pigmentation of the iris continues until adolescence, and varies little after that. Research by Worrall [40] and Berggren [41] discard the claims that iris texture varies with changes in health (a theory often referred as *iridology*). Flom and Safir’s patent [21] describes an automatic iris identification system, and is often regarded as a significant contribution in the area of iris biometrics. Davson’s work [30] on the physiology of the eye describes that the phenotypic expressions even of two irises with same genetic genotype (e.g., identical twins, or the left and right pair possessed by one individual) are uncorrelated. Newell’s work [42] confirms that after adolescence, a healthy iris varies little for the rest of a person’s life.

The first prototype iris recognition system was described by Johnson [43]. Later in 1993, Daugman’s landmark paper [33] described a method to perform rapid, automatic iris recognition. Johnston and Grace [44] list the benefits of using iris, over other biometric traits, for identification of individuals. A recognition performance analysis by Bouchier et al. [45] suggests that iris can serve as a high-confidence biometric identifier.

1.2.4 Steps involved in Iris Recognition

The function of an automatic iris recognition system is to extract, represent, and compare the textural intricacy of the surface of the iris. The key steps involved in a generic iris recognition system can be listed as follows:

1. **Image acquisition:** The input raw data (images or videos) is usually acquired using a sensor of adequate resolution to capture the iris texture. Such data is either processed immediately, or stored in a database for later processing. The usefulness of an acquired iris image mainly depends on the quality, and the spatial extent of the iris present in the captured image. Both these factors can be regulated at the image acquisition stage to achieve reliable accuracy. Most iris recognition systems require a considerable level of user cooperation. Current research is progressing towards successful acquisition of iris images in less constrained environments [9].

2. **Iris segmentation:** Depending on the field of view of the sensor, the acquired images can potentially include other regions of the eye (e.g., the sclera, eyelashes, eyelid, eyebrow, surrounding skin regions, etc.). Given such image, the annular region encompassed between the pupillary and limbic boundaries has to be identified to perform feature extraction. Furthermore, anatomical features that usually occlude the iris texture (e.g., eyelashes and eyelids) have to be excluded from the consideration. This process of automatically locating the iris boundaries, and excluding the noisy regions is called as iris segmentation. This step is often considered critical, as an incorrect segmentation can negatively impact the iris recognition performance [46].
3. **Normalization:** The size of an iris can vary significantly due to its dilation, contraction, resolution of the sensor and the imaging distance [47]. To address such variations in size, the segmented iris is usually unwrapped to a normalized coordinate system. This normalization operation is performed by representing the segmented iris as a rectangular image, the rows of which correspond to the concentric regions of the unsegmented iris. A widely popular technique for iris normalization, Daugman's *rubber sheet model* [48], re-maps every point in the segmented iris region to a pair of polar coordinates. While it is a popularly used step before performing matching, certain approaches exist that do not require iris normalization. Such approaches are referred as *segmentation-free approaches* [49].
4. **Feature extraction and matching:** Feature extraction refers to the process of encoding the discriminatory information obtained from the segmented (or, normalized) iris, into a feature vector. In the matching stage, the feature vector obtained from a probe image is compared with the other feature vectors in the gallery to perform recognition. Various techniques have been proposed in literature [47] to perform feature extraction and matching. The focus of all such techniques is to reduce the computational time and complexity, while improving the recognition performance.

The basic modules of an iris recognition system are shown in Figure 1.5.

Within the past decade, a wide variety of research has emerged in the field of iris recognition [47]. While very few improvements have been suggested towards image acquisition [9, 50], a majority of the research has been focussed on improving the segmentation,

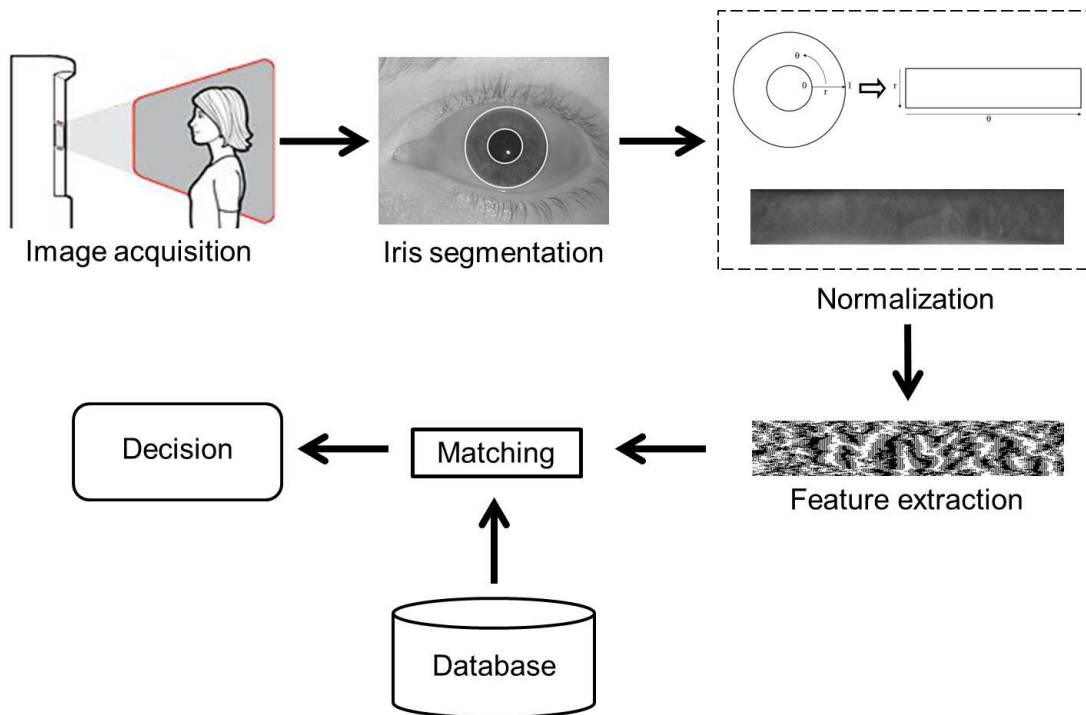


Figure 1.5: A block diagram of an iris recognition system.

encoding and matching schemes.

1.3 Existing Iris Recognition Research

1.3.1 Image Acquisition

Various types of image sensors have been used for iris image acquisition. Based on their portability and the convenience offered to the users, iris sensors can be divided into the following categories:

1. *Portable, or hand-held sensors*: These sensors offer portability, and can be designed to be functional in rugged situations. However, such sensors could require effort from two individuals (subject, and an operator) to acquire a good quality iris image. An operator is required to align the sensor with the subject's eye. Examples of this type of sensors include Datastrip Easy Verify and Retica Mobile-Eyes.
2. *Fixed, or wall-mounted sensors*: These sensors are some of the most widely used, and can be deployed in both covert and overt situations. Such sensors can eliminate the requirement of an additional operator during image acquisition (given the subject is

acquainted with the imaging system). Examples of this type of sensors include LG IrisAccess 4000 and Oki IrisPass.

3. *Portals, or walk-through sensors*: These sensors are typically used in situations involving rapid iris recognition for a large volume of users (e.g., airports). Images of the iris are captured when the user passes through a portal like structure. The illumination sources and image sensors are usually mounted on the walls of the portal. An example of such imaging system includes Iris on the Move Passport setup [9].

A sample image from each type of sensor is provided in Figure 1.6.



Figure 1.6: Examples of the three different types of iris sensors: (a) Retica Mobile-Eyes, (b) Oki IrisPass and (c) Iris on the Move (IOM) Passport system.

A detailed description of a traditional iris image acquisition system was first provided by Wildes [51]. Negin et al. [16] proposed an image acquisition system that partially relaxes the requirement of controlled user interaction. Matey et al. [9] describe a walk-through image acquisition system that provides complete relaxation of the constraints on the users. Recent improvements in iris image acquisition include the usage of wavefront coding [52,53], and hyper-spectral imaging systems [54].

1.3.2 Imaging Wavelengths

Iris images are typically acquired using sensors that operate in the near-infrared (NIR) spectrum. The wavelength of the illuminating sources range between 700-900nm. The usage of NIR spectrum for iris recognition provides two critical benefits:

1. It is observed that the effect of melanin, a color inducing compound, is negligible at longer wavelengths. Using NIR spectrum ensures that the acquired image reveals information related to the texture of the iris, rather than its pigmentation.
2. Compared to the visible spectrum, the texture of dark colored irides can be well observed using NIR.

Despite the benefits offered by the NIR spectrum, it is critically important to study iris recognition using images acquired under the visible spectrum. Some of the reasons to support such research can be listed as follows:

1. Most surveillance cameras work in visible wavelength. When iris recognition has to be performed on the images acquired using such systems, the iris data would be in visible wavelength.
2. The current sensor and illuminator technology makes it difficult to acquire iris images from a distance using sensors that operate in the near infra-red wavelength.
3. Iris images acquired under visible spectrum reveal a high level of information related to its pigmentation. Such information can be used for iris classification purposes.

Considering the above listed benefits, iris recognition has been studied under near-infrared, visible, and multi-spectral wavelengths. Boyce et al. [55] performed a comparison of

iris recognition performance under visible and near-infra red spectrum. A study of the iris recognition performance using NIR wavelengths ranging between 900-1450nm was carried out by Ross et al. [56]. More recently, Ives et al. [57] discussed the performance of iris recognition using illumination sources of wavelengths ranging between 405-1070nm.

1.3.3 Iris Image Quality

The quality of an iris image is a critical component in obtaining a high recognition performance. Iris image quality can be assessed at two different stages, with each stage offering a unique benefit:

- *Image acquisition stage*: Allows the system to prompt for re-acquisition, if the image quality is poor.
- *Matching stage*: Helps in weighting the match score with the quality score, to minimize the impact of a poor quality image on the matching performance.

Chen et al. use 2D wavelets to assess the quality of a localized iris image [58]. Kalka et al. [59, 60] compute the quality of an iris image by examining seven different factors: defocus blur, motion blur, deviated gaze, occlusions, lighting, specular reflection, and pixel counts. A likelihood ratio based fusion scheme to combine the quality scores obtained by evaluating the defocus, motion blur, and deviated gaze was used by Li et al. [61].

1.3.4 Iris Segmentation

Typically, the variation in image intensities across the pupillary boundary is much stronger than the variation across the limbus boundary. This property lends to the fact that, in most cases, the pupillary boundary can be detected using simple intensity thresholding operation. On the other hand, determining the limbus boundary can be a comparatively difficult task. One of the earliest works on iris segmentation was performed by Daugman [62]. An integro-differential operator is used to detect the iris boundaries, that are approximated as perfect circles. Given an iris image $I(x, y)$, it is first convolved with an image smoothing function (e.g., a Gaussian filter). The smoothening step helps in (a) attenuating the effect of noise (e.g., sensor noise) in the image, and (b) eliminating undesired weak edges (e.g., boundaries within the iris), while retaining the desired strong edges

(e.g., iris boundaries, eyelid boundaries, etc.). An integro-differential operator (IDO) is then used to search for the maximum value of a normalized integral along circular contours of varying radii and center coordinates. The search process over the image domain (x, y) , using an IDO, can be mathematically expressed as:

$$\max(r, x_0, y_0) \left| G_\sigma(r) * \frac{\partial}{\partial r} \oint_{r, x_0, y_0} \frac{I(x, y)}{2\pi r} ds \right|, \quad (1.1)$$

where

$$G_\sigma(r) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(r-r_0)^2}{2\sigma^2}\right) \quad (1.2)$$

represents the radial Gaussian with a center r_0 and standard deviation (scale) σ , which is used for image smoothing. The symbol $*$ denotes the convolution operation, and r represents the radius of the circular arc ds , centered at the location (x_0, y_0) . The division by a factor of $2\pi r$ normalizes the circular integral with respect to its perimeter. In other words, the IDO behaves as a circular edge detector, that searches iteratively for the maximum response of a contour path defined by the parameters (x_0, y_0, r) . Depending on the values of the radii considered, the optimal parameters of the IDO are treated as either the pupillary or limbus boundaries. Once both the iris boundaries are detected, the boundaries of the eyelids can be detected by changing the integration path of the operator from circular to arcuate. Figure A.2 shows the output of an IDO when used to detect the iris and eyelid boundaries.

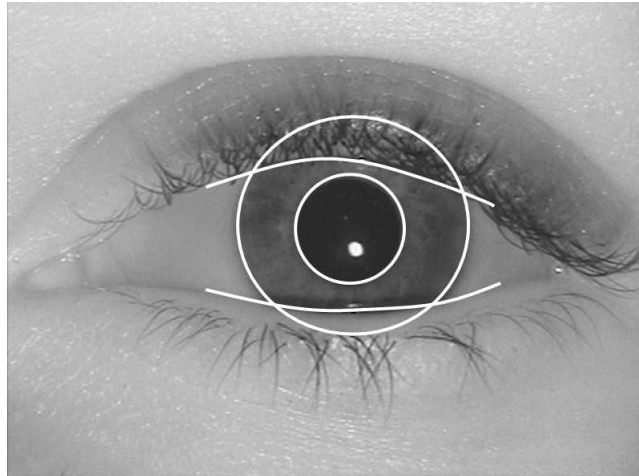


Figure 1.7: Output obtained by applying an integro-differential operator to detect both the iris and eyebrow boundaries.

Another widely popular approach for performing iris segmentation [51] uses Hough

transforms. Similar to IDO, this technique also approximates iris boundaries as perfect circles, and the eyelid boundaries ellipses. A wide number of iris segmentation approaches improve on the idea of Hough transform [63–66].

The approximation of the iris boundaries as perfect circles can be accepted when an iris image is acquired under near-ideal conditions from a cooperative subject. In an image acquired under non-ideal conditions, the limbus boundary may not be completely circular (due to the occlusions, gaze deviations, etc.). Some of the significant approaches that can perform segmentation of non-circular irides use Geodesic Active Contours (GAC) [67], variational level sets [14], Fourier-based approximations [62], Active Shape Models (ASM) [68], graph cuts [69], iterative directional ray based segmentation [70]. Some of the segmentation free approaches include the usage of SIFT features [49, 71]. A fairly detailed survey of iris segmentation approaches is presented in [72, 73]. Research on improving iris segmentation includes techniques for eyelash detection and removal [17, 64, 74], segmentation error prediction [75], modifications [76] to existing open-source segmentation implementations [77] and segmentation techniques for images acquired under visible wavelength [78–80].

1.3.5 Iris Encoding and Matching

Several iris encoding algorithms have been proposed in the literature. One of the most widely used scheme was proposed by Daugman [62], that uses a multi-scale 2D Gabor wavelet transforms to encode a normalized iris. Given a normalized iris image in polar coordinates, a 2D Gabor wavelet (in polar coordinates) can be expressed as:

$$H(r, \theta) = e^{-i\omega(\theta-\theta_0)} e^{-(r-r_0)^2/\alpha^2} e^{-i(\theta-\theta_0)^2/\beta^2}, \quad (1.3)$$

where (r_0, θ_0) denote the center frequency, (α, β) denote the effective width and length, and ω denotes the spatial frequency of the wavelet. The output of the Gabor wavelets is then demodulated to compress the data into a feature vector. This is performed by quantizing the phase information into four different levels, one for each quadrant of the complex plane. The modulation and phase quantization process can be mathematically represented by the following equation:

$$h_{Re, Im} = \text{sign}_{Re, Im} \int_r \int_\theta I(r, \theta) e^{-i\omega(\theta-\theta_0)} e^{-(r-r_0)^2/\alpha^2} e^{-i(\theta-\theta_0)^2/\beta^2} r dr d\theta, \quad (1.4)$$

where $h_{Re,Im}$ is the complex valued bit whose real and imaginary components are dependent on the sign of the integral. The response of this operation is a binary output, usually called an *IrisCode*. The normalized Hamming Distance (HD) between two IrisCodes is used as a measure of dissimilarity between two irides. This value is computed by masking every IrisCode with its respective mask, to disregard the noisy regions. The Hamming Distance between two IrisCodes is computed by:

$$HD = \frac{\| (IrisCodeA \oplus IrisCodeB) \cap MaskA \cap MaskB \|}{\| MaskA \cap MaskB \|} \quad (1.5)$$

The XOR operator detects the bits that disagree between the two IrisCodes, while the AND operator masks the noisy regions. The denominator helps in normalizing the total number of bits that disagree to a value between $[0, 1]$. A perfect match between two IrisCodes would result in a HD of 0.

Another popular approach to encode iris include the usage of a Laplacian-of-a-Gaussian (LOG) filter [51]. Matching is performed using the normalized correlation between the test and training images. Other encoding approaches use multi resolution Independent Component Analysis (ICA) [81], PCA and ICA [82], ordinal features [83, 84], correlation filters [65], 2D phase congruency method [85], local key variations of the iris [64], and SIFT (after normalization [86], and without normalization [49]).

1.3.6 Iris Recognition under Near-Ideal Conditions

The iris is considered to be a robust and unique biometric with a very low False Accept Rate (FAR). Daugman [87] suggests that the iris pattern has almost 250 independent degrees-of-freedom of textural variation. Very high recognition performances can be obtained using iris images acquired under controlled conditions, with cooperative subjects. This claim is supported by some large scale performance evaluations such as Daugman's observations from 200 billion iris pair comparisons [48] and NIST Iris Exchange (IREX) evaluations [88].

1.4 Ocular Recognition

Depending on their field-of-view, most iris sensors capture the surrounding region of the eye, without additional requirements. As mentioned in Section 1.1, such ocular region

can serve as a soft biometric trait and could be used to improve the overall recognition performance. Some of the anatomical characteristics and features that contribute to the uniqueness of ocular region are shown in Figure 1.8.

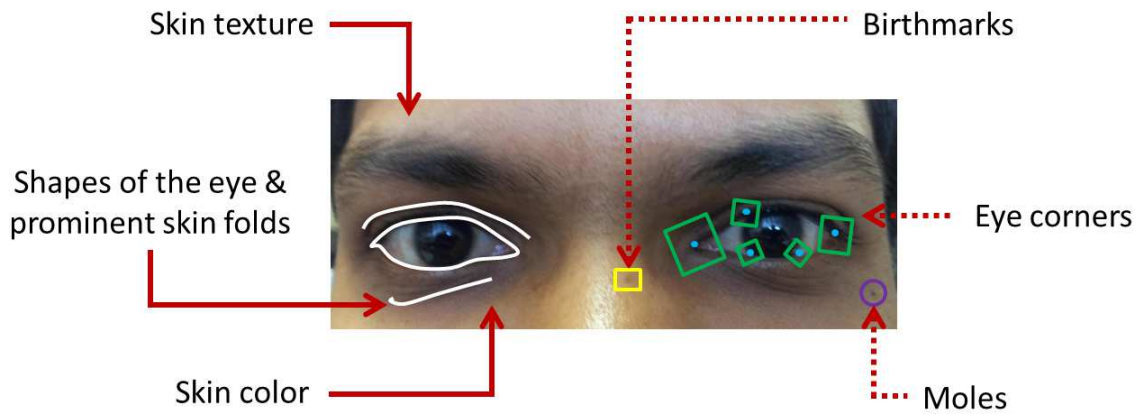


Figure 1.8: Anatomical characteristics and features that contribute to the uniqueness of ocular region.

The benefits of using ocular region for recognition can be listed as follows:

1. *Minimal imaging requirements*: The designing, or requirement, of a newer sensor is eliminated as the ocular regions can be easily captured using an iris or face sensor.
2. *Distance trade-off*: The ocular region represents a good trade-off between using the entire face region or using only the iris texture for recognition. When the entire face is imaged from a distance, the iris information could be of low resolution. On the other hand, when the iris is imaged very closely, the recognition system is forced to rely only on the iris. However, the ocular biometric can be useful over a wide range of distances.
3. *Supplementary information*: The ocular region can offer information about the eye shape, skin color, etc., that could be used to further improve the overall recognition performance [89].
4. *Challenging conditions*: Ocular information can be of significant importance in challenging conditions involving:
 - (a) Occlusions: Face recognition performance can be negatively impacted due to occlusions (e.g., presence of scarves, facial hair, etc.). Similarly, iris recognition

performance is also reduced when the iris is occluded due to deviated gaze, eyeglasses or blinking eyelids. In such cases, the ocular information could be reliably used (see Figure 1.9).

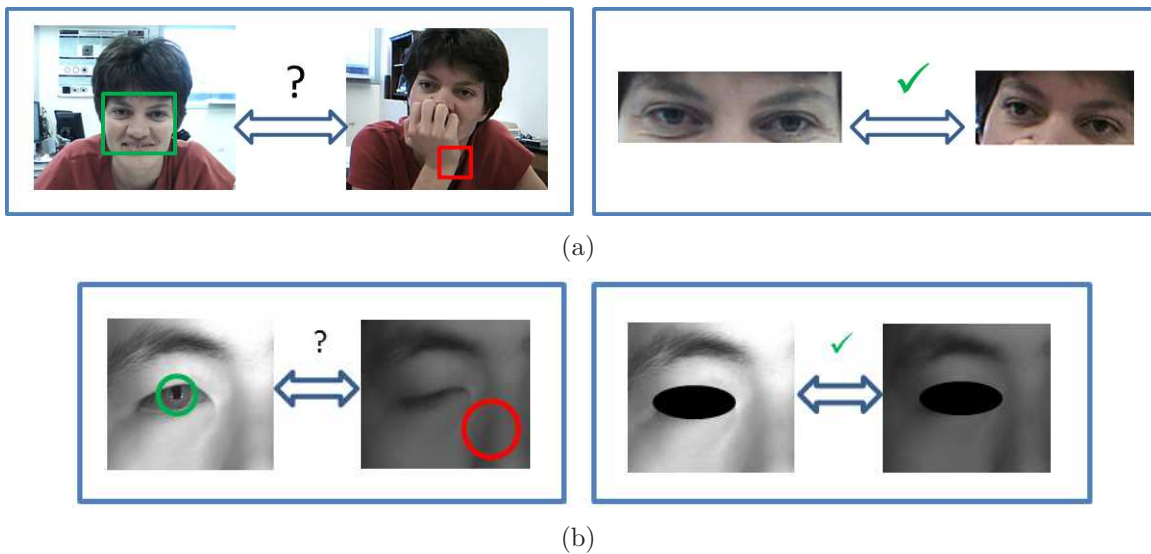


Figure 1.9: Sample situations where ocular information can be used to improve (a) face, and (b) iris recognition performance, respectively.

(b) Spoofing or alterations: In cases where an iris is spoofed using, for example, texture imprinted on a contact lens, the ocular information could be used to confirm or refute an identity. Similarly, ocular information can be used improve the recognition performance when the facial appearance is altered due to plastic surgery (this is explained later, in Chapter 3).

5. *Privacy concerns*: From a social perspective, ocular region could potentially be used to resolve some biometric privacy concerns. Unlike face recognition, ocular region information could be obtained without violating some societal considerations (see Figure 1.10).

Ocular recognition may not be robust to some variations:

- **Expressions**: Variation in human expression (smiling, neutral, etc.) can cause non-linear deformation around the eye, thereby reducing the ocular recognition performance.
- **Aging**: Under a larger time frame consideration, wrinkles and folds around the eye caused by aging could change the overall appearance of the ocular region.



Figure 1.10: Ocular information could be acquired with minimal privacy violation.

1.4.1 Existing Ocular Recognition Research

Park et al. [90] first introduced the concept of ocular recognition using images acquired in the visible spectrum. The recognition process consists of: localizing the ocular region, image alignment, feature extraction, and matching. Typically, the iris location and radius information is used for localization, alignment and scale normalization of the ocular region. Many feature extraction techniques have been proposed in the existing literature to perform ocular recognition. Depending on the region of interest from which the features are extracted, a majority of the existing techniques can be classified into two categories: global (e.g., GOH [90], GIST [91], etc.) or local (e.g., SIFT [90,92]). While global feature extraction techniques summarize features from the entire image (e.g., shape, color, texture, etc.), local feature extraction techniques gather information around a set of detected key points. Examples of images showing the sampling point patterns for global and local feature extraction schemes are shown in Figure 1.11. The extracted information is summarized using histograms, and matched by computing the corresponding Euclidean distances. Existing studies report reasonably good ocular recognition performance when using just one type of feature extraction scheme. Such a performance can be mainly attributed to the high quality of the input images.

Currently, less research exists in the area of ocular recognition as it is a relatively new topic in biometrics. Woodard et al. [23] and Bhatt et al. [91] show that ocular information can aid iris recognition under non-ideal image acquisition conditions. Work by Park et al. [24] suggests that ocular information can be used in situations where face recognition may fail (e.g., facial occlusions caused by scarves). It has also been observed that the ocular information can provide a reasonable gender identification performance [93]. Studies by Hollingsworth et al. [94,95] provide a benchmark for comparing human and machine

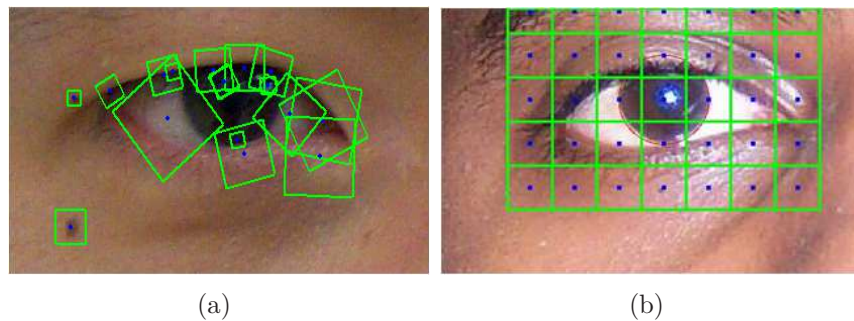


Figure 1.11: Sampling point patterns for (a) local (e.g., SIFT) and (b) global (e.g., GOH) ocular feature extraction schemes.

performance using ocular information. More recent studies have indicated that the ocular region can provide significantly higher performance than the iris, when imaged under highly non-ideal conditions [4, 96].

1.5 Factors that affect Iris and Ocular Recognition

Various factors that can render iris and ocular recognition challenging, are listed as follows:

1. *Poor illumination*: An iris image acquired under poor illumination may not reveal the richness of the iris texture. Furthermore, performing iris segmentation would be extremely difficult, as the image may offer minimal or no information about the iris boundaries. Similarly, an ocular image acquired under low illumination may not offer enough information regarding the iris radius or location, thereby rendering ocular recognition very challenging. Figure 1.12 shows an iris, and an ocular image acquired under poor illumination conditions.
2. *Specular reflections*: Specular reflections are small regions in an iris image characterized by pixels of high intensity values, that are typically caused by improper focusing of the light source. If specular reflections are present on (or even close to) the iris boundaries, iris segmentation becomes difficult. Specular reflections that overlap with the iris texture can induce noise, thereby lowering the recognition performance. Figure 1.15 shows an iris image with specular reflections.
3. *Stand-off distance*: Stand-off distance refers to the distance of the camera from the

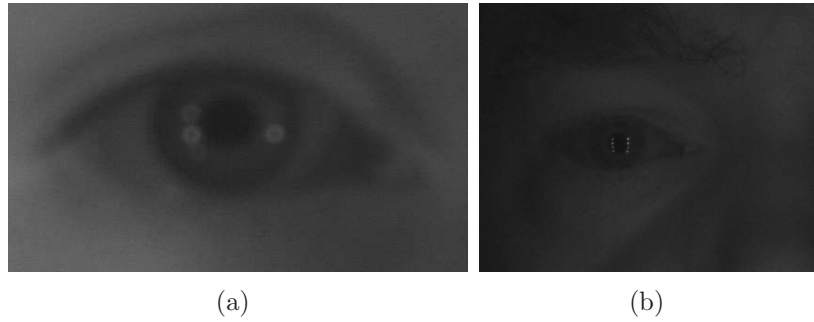


Figure 1.12: (a) Iris, and (b) ocular image acquired under poor illumination conditions. Although the iris boundaries or location can be determined by humans, automatic recognition using such images is extremely difficult [4].

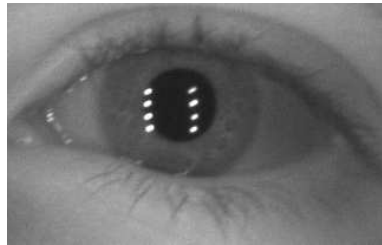


Figure 1.13: An iris image containing specular reflections on the pupillary boundary.

subject. If the stand-off distance is large, the resolution (number of pixels occupied by the iris region in an image) can be low. In such cases, the textural richness of the iris observed in the image could be reduced, leading to a lower recognition performance. Figure 1.14 shows an iris image acquired at a large stand-off distance.



Figure 1.14: Closeup of an iris image acquired at a large stand-off distance.

4. *Image sensors*: The following factors related to image sensors play a significant role in acquiring a good quality iris or ocular image:

- (a) Type of sensor: Iris sensors that acquire images in the Near-Infra Red (NIR)

spectrum are preferred over sensors that operate in the visible spectrum. This is based on the fact that even dark colored irides can reveal detailed texture information when imaged in the NIR spectrum.

- (b) Resolution of the sensor: A high resolution sensor that can capture iris images with a minimum diameter of 200 pixels are preferred over other low resolution sensors.
- (c) Positioning of the sensor: The positioning of the sensor plays an important role in acquiring an iris image of good quality. For example, if the iris sensor is placed above or below the eye level of a subject, the iris region may not be captured in its entirety.
- (d) Sensor noise: Although not seen as a major factor, sensor noise can produce artifacts in an image, thereby affecting iris and ocular recognition. Sample iris and ocular images containing sensor noise are shown in Figure 1.15.

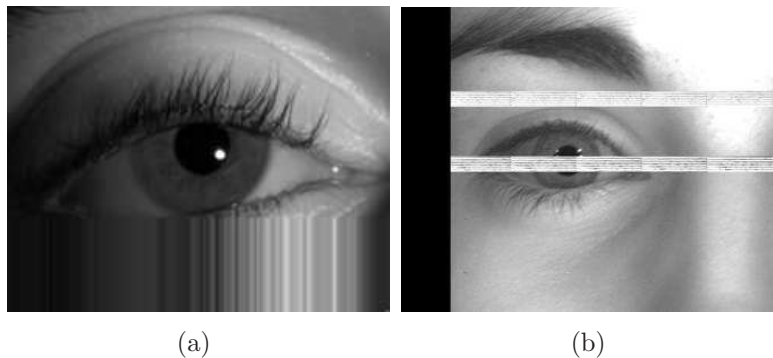


Figure 1.15: Sensor noise in an (a) iris and (b) ocular image.

5. *Eyelids*: Eyelids are thin folds of skin that cover and protect the eye from foreign bodies and extreme lighting. The movement of eyelids can be both voluntary (e.g., closing eyelids when tired), or involuntary (e.g., blink caused by a reflex). To obtain an un-occluded image of the iris, the user is required to hold the eyelids wide open for a brief period of time during image acquisition. However, under normal conditions, a minor portion of the human eye is typically occluded on the top and the bottom, by the upper and the lower eyelid, respectively. Such occlusions can reduce the area of iris image that can be captured in an image. Figure 1.16 shows an iris image exhibiting eyelid occlusion.

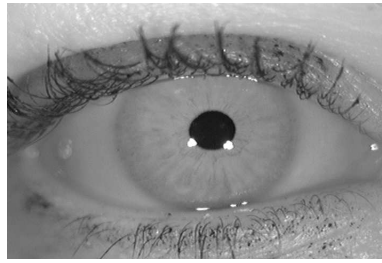


Figure 1.16: An iris image showing occlusions caused by the eyelids.

6. *Eyelashes*: Eyelashes are the hair at the end of the eyelids. Like eyelids, eyelashes also provide protection to the eye from external debris. Although the occlusions caused by eyelashes are minimal, they can impact the process of iris segmentation. This is due to the fact that eyelashes can cause uneven interruptions at the limbus boundary. Empirical observations reveal that eyelash occlusion is typically more pronounced in Asian subjects, due to the presence of the epicanthic fold. Figure 1.17 shows an iris image with eyelash occlusions.

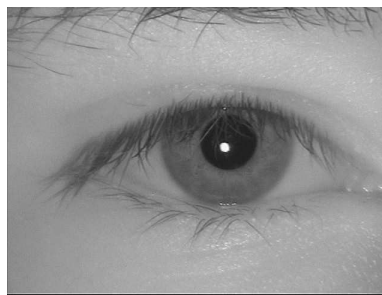


Figure 1.17: An iris image showing occlusions due to eyelashes.

7. *Nature of the interacting population*: To acquire a good quality iris or ocular image, it is required for the target population to be cooperative and habituated with the image acquisition system.
8. *Outliers*: In rare cases, diseases and abnormalities of the iris can impact the segmentation and recognition performance (e.g., congenital abnormalities). A sample image of such case is shown in Figure 1.18.
9. *Eye glasses or contact lenses*: If a user wears eye glasses or contact lens, the acquired iris or ocular images may suffer from additional reflection artifacts due to

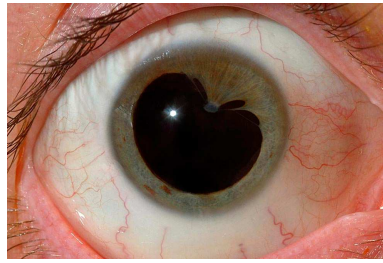


Figure 1.18: Iris image of a subject suffering from congenital iris abnormality. Notice that the pupil boundary is neither circular nor elliptical. Image taken from [5].

these entities. In recognition system involving cooperative subjects, this problem can be minimized by requesting the user to avoid wearing eye glasses during image acquisition. However, if a subject wears contact lenses (cosmetic/non-cosmetic), it may not be convenient for the user to remove them, even in a cooperative image acquisition setup. Research has shown that contact lenses can impact the performance of iris segmentation and recognition [6]. Figure 1.19 shows iris and ocular images of two different users wearing contact lens (non-cosmetic, hard lens) and eye glasses, respectively.

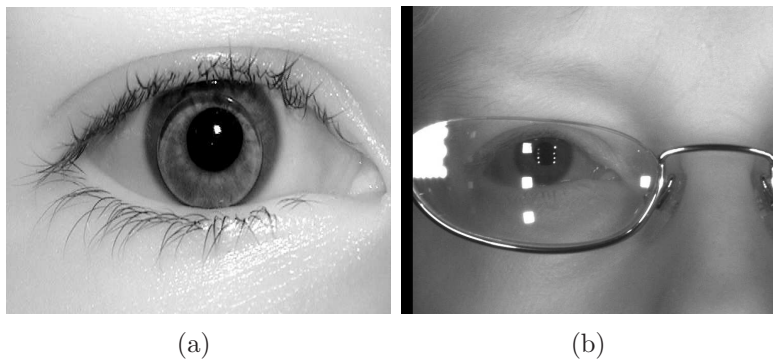


Figure 1.19: (a) An iris image of a subject wearing contact lens. (b) Ocular image of a subject wearing eye glasses. Image in (a) taken from [6] ©Elsevier.

10. *Motion blur*: Motion blur in iris and ocular images can occur mainly due to three reasons: (a) when the image is acquired from a moving subject, (b) movement of the camera, and (c) movement of the subject's eye while adjusting to the device and the environment. In iris images containing motion blur, the texture of the iris is blurred, thereby impacting the iris encoding. On the other hand, motion blur in ocular images can impact the feature encoding, thereby reducing the recognition

performance. Figure 1.20 shows iris and ocular images containing motion blur.



Figure 1.20: An (a) iris and (b) ocular image containing motion blur.

11. *Deviated gaze*: Iris images containing deviated gaze are observed when the sensor is not orthogonal to the plane of the iris. In such cases, the surface area of the iris region captured is reduced than normal. Deviated gaze iris images are typically caused in situations where the image acquisition is non-ideal, i.e., when the subject is in motion, or not aware of the image acquisition process. Figure 1.21 shows an example of a deviated gaze iris image.

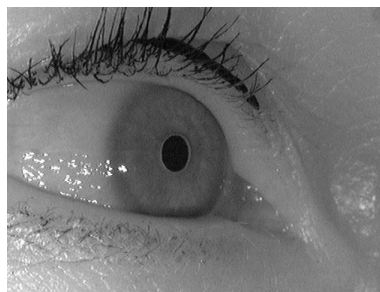


Figure 1.21: An off-angled iris image.

12. *Aging*: Flom and Safir [21] have postulated that the basic, significant features of the iris remain extremely stable and do not change over a period of many years. However, the claim regarding the permanence of the iris has been challenged in the recent literature [6]. This effect is also shown to slightly reduce the recognition performance. Similarly, aging can cause significant variations within the ocular regions. While research in this area is still pending, a drop in ocular recognition performance could be expected due to the variation in the overall appearance.

13. *Deformation*: The ocular region appearance can significantly vary due to the deformations occurring around the eye. Such deformations are non-linear, non-rigid, and are quite difficult to model using the existing techniques. Ocular recognition using such images is a challenging task. Figure 1.22 shows the variation in the appearance of ocular region due to deformations, caused by eye-blinking.

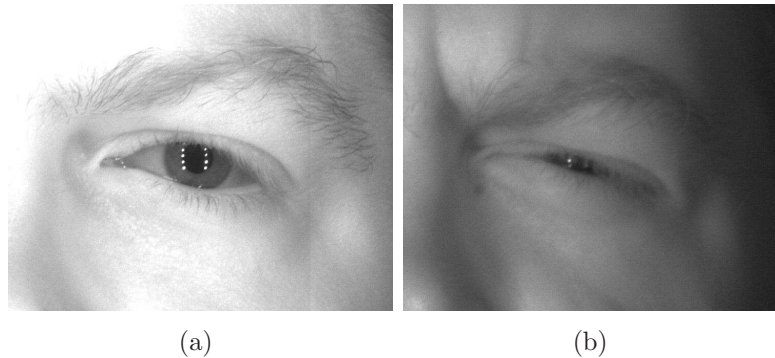


Figure 1.22: Images showing the variation in appearance of ocular region: (a) without, and (b) with deformations, respectively. Image source: 04233d1632-08-l.jpg and 04233d1649-03-l.jpg from [7].

14. *Image compression*: Under practical considerations involving large number of users, iris images could be compressed before storing in a database to reduce the storage requirements. Daugman [97] suggests that the recognition performance can be impacted by such image compression process. Several recommendations have been proposed to limit the image compression factor for iris images [88]. An example of a compressed iris image is provided in Figure 1.23.

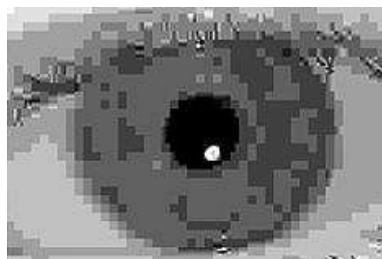


Figure 1.23: An iris image with artifacts caused by severe image compression. Image taken from [8].

1.6 Motivation

Given good quality iris or ocular images (i.e., images that do not exhibit the non-ideal factors listed in Section 1.5), the recognition performance can be expected to be nearly perfect. However, obtaining such images is a very controlled process. The users are typically required to be stationary (or walking at a very slow pace), and maintain a fixed gaze at a specified location for a short period of time. Such constraints have generated an increasing amount of interest to advance ocular biometrics under challenging conditions. The most invested effort in this regard is towards minimizing the user constraints, while improving the recognition performance. Following is a list of sample scenarios that render ocular recognition challenging, thereby *motivating* the need for extensive research in this area:

1. *Recognition at longer distances*: When performing ocular recognition involving large stand-off distances, the quality of the captured image could be significantly reduced (even with cooperative subjects). An example of such scenario, as shown in Figure 1.24(a), could be an access point involving subjects walking through a portal.
2. *Surveillance under covert conditions*: In some cases, surveillance could be carried out under covert conditions where the subjects are unaware of the recognition process. Therefore, a large number of images could suffer from a combination of non-ideal factors, thereby negatively impacting the recognition performance. Figure 1.24(b) shows a practical example of such situation involving watch-list screening at an access point.
3. *Limited control over image acquisition process*: Three major factors contribute to the quality of an image acquired during image acquisition: (a) cooperation of the users, (b) ambient conditions, and (c) sensor limitations. While it is desired to optimize all the factors during image acquisition, such control may not be practically feasible in every situation. Example scenarios of this consideration include availability of only a particular type of sensors (e.g., visible), involvement of users that are unacquainted with the system, and poor ambient conditions (e.g., low illumination).
4. *Cross-source matching*: In many practical applications, it is possible to encounter the

problem of matching ocular images obtained from different sources. Some examples of this consideration include:

- *Cross-sensor matching* - Matching scenarios in which the sensor used for acquiring the probe images is different from that used during enrollment. The probe and gallery images could exhibit variations in resolution, image formats and sensor parameters (e.g., depth of field, field of view, illumination source, etc.), thereby impacting the recognition performance. A sample pair of ocular images acquired using two different sensors are shown in Figure 1.24(c).
- *Cross-modality matching* - Matching scenarios involving probe and gallery images that differ with regards to the biometric modality captured (e.g., face, iris, and ocular). Such scenarios could arise when the availability of the iris images is limited (e.g., forensic application where the iris region in a face image acquired using a surveillance camera, has to be matched with the iris images stored in a database). Cross-modality ocular recognition can present significant challenges during the automatic localization, feature extraction and matching stages. An example of cross-modality matching is shown in Figure 1.24(d).

5. *Spoofing*: An iris recognition system is vulnerable to a spoof attacks if a fake iris sample (e.g., contact lens printed with artificial patterns) is presented. Detection of spoofed iris images is a significant challenge, and is currently an active research topic. As shown in Figure 1.24(e), using the additional information from the surrounding ocular information could potentially remedy this problem.

1.7 Contributions

The main scope of this thesis is in improving performance of iris and ocular recognition systems under unconstrained conditions. The major contributions of this thesis are:

- **Improving low-resolution iris recognition**: A technique for improving iris recognition involving low-resolution images is proposed. The proposed approach performs image-level using Principal Components Transform (PCT) and image averaging. The

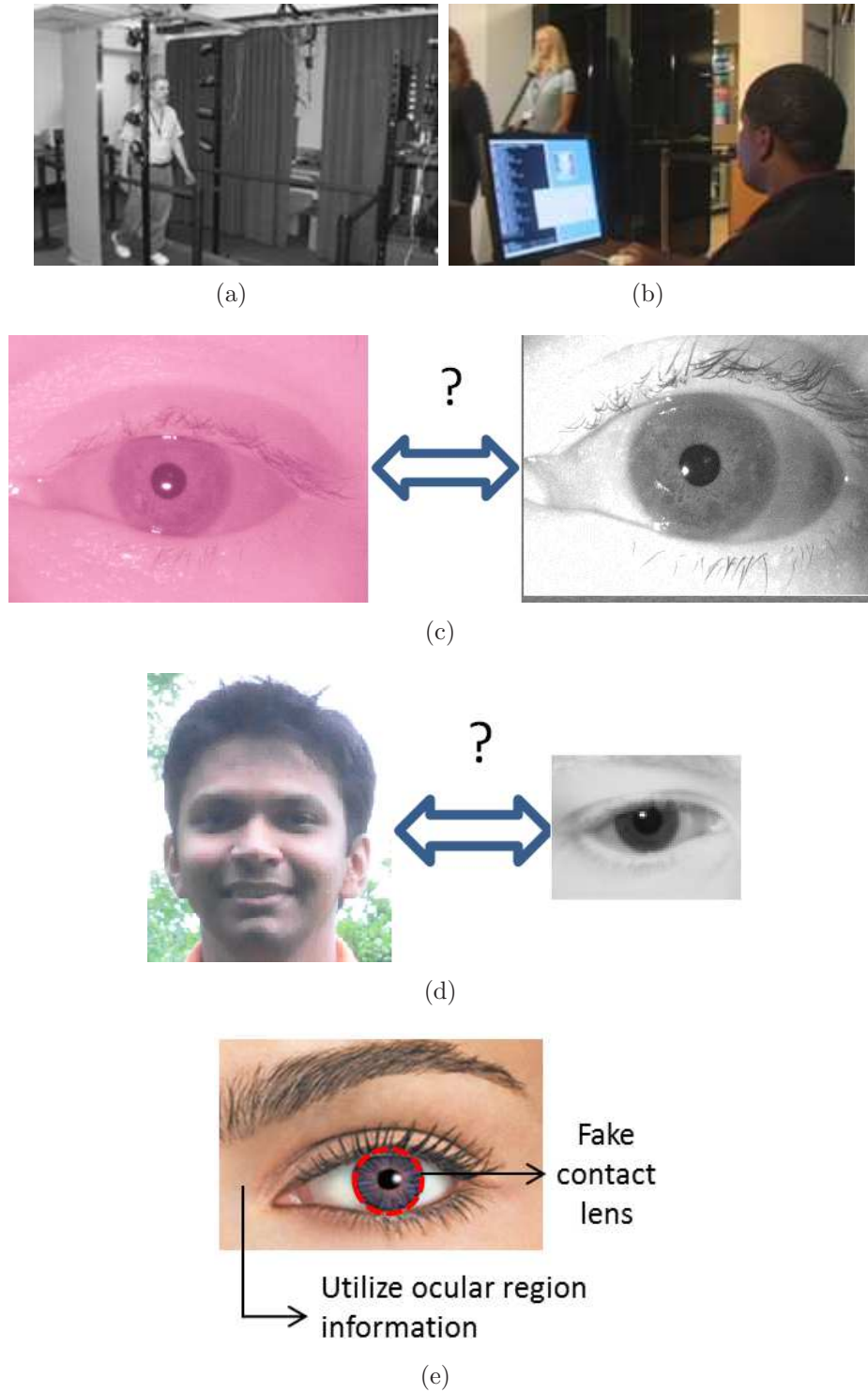


Figure 1.24: Sample scenarios that motivate the need for rigorous research in ocular recognition: (a) recognition at a distance (image from [9]), (b) covert surveillance (image from [10]), (c) cross-sensor matching (image from [11]), (d) cross-source matching (author's personal images), and (e) spoofing (image from [12]).

proposed approach is observed to outperform the existing image-level fusion schemes related to this problem.

- **Mitigating effects of plastic surgery using ocular information:** The feasibility of improving overall recognition performance in plastic surgery images by combining face and ocular information is demonstrated. It is shown that such fusion leads to improved recognition performance, when compared to using face recognition only.
- **Matching ocular regions in face-iris image pairs:** The advantage of using ocular region in matching *RGB* face and *NIR* iris images is demonstrated. A sparse representation approach is proposed which generates a joint dictionary from corresponding pairs of face and iris images. The proposed approach is observed to outperform existing ocular recognition techniques. Additionally, it is observed that ocular region provides better recognition performance when compared to iris, in the considered challenging database.

1.8 Thesis Organization

This thesis focusses on improving iris and ocular recognition under non-ideal conditions. Chapter 2 describes a technique for improving iris recognition performance in low-resolution images. Chapter 3 deals with improving face recognition performance after plastic surgery, using ocular region information. Chapter 4 describes the ocular matching problem in *RGB* face and *NIR* iris image pairs.

Chapter 2

Low Resolution Iris Recognition

2.1 Introduction

There has been a steady increase in attention towards achieving reliable iris recognition from a distance. However, a good recognition performance can be guaranteed only when cooperative subjects, within a short distance from the sensor, are involved. Under unconstrained conditions, the images may be of poor quality, thereby reducing the recognition performance. The present chapter discusses iris recognition involving low-resolution imagery.

A wide number of factors can impact the resolution of an iris image. Some of such factors include: (a) stand-off distance (i.e., distance of the subject from the sensor), (b) resolution of the sensor, (c) ambient conditions (e.g., atmospheric turbulence), and (d) moving subjects in an unconstrained image acquisition environment. While the effects of other factors can be mitigated by using a high resolution sensor, it is difficult to handle the stand-off distance problem. An increase in the stand-off distance can reduce the size (or pixel resolution) of the eye recorded in an image, when a fixed zoom-factor sensor is used. Such a resolution reduction can lower the textural quality of the iris in the image, which can in turn affect the performance of the recognition system. Figure 2.1 illustrates this effect.

When the input images are of poor quality, fusion methods can be used to enhance the recognition performance. Biometric fusion refers to the process of aggregating the information needed for reliable recognition from multiple sources of evidence [98]. Depending on the type of information available, fusion can be performed at various levels in a biometric system (e.g., image-level, feature-level, score-level, etc.). In this chapter, an image-level

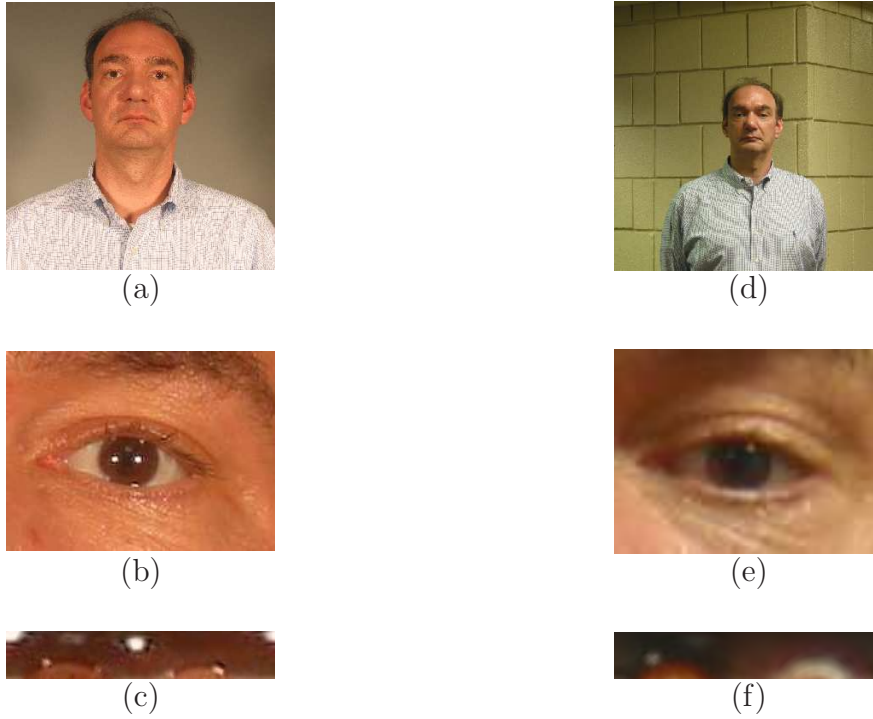


Figure 2.1: Panels (a), (b), and (c), respectively, show the face, the eye, and the normalized iris regions (in the visible spectrum) of a subject standing close to the sensor. The corresponding regions for a subject with larger stand-off distance are shown in panels (d), (e), and (f), respectively.

fusion scheme is presented that uses the information contained in the multiple frames of an iris video. The use of multi-frame iris fusion has several benefits:

1. In many image-level fusion techniques, registration (or alignment) of the input images into a single coordinate system of reference is very important. Registration of images obtained at different time instances or from different sensors, is a challenging task. However, the frames extracted from a given iris video are likely to be aligned. As a result, the errors caused by improper image registration can be greatly reduced.
2. The frames within an iris video contain information related to the spatio-temporal activity of the iris and its surrounding region over a short period of time. As this information is continuous, good quality frames can be selectively chosen for fusion while avoiding poor quality frames.

The proposed technique performs image-level fusion in two stages: (i) by first applying Principal Components Transform (PCT) to the individual frames, and (ii) then averaging

the resulting images. The performance of image-level fusion is compared against that of score-level fusion. Experimental results, in both cases, indicate that the fused outputs provide better recognition performance than their corresponding low-resolution source images.

2.2 Image-level Fusion

2.2.1 Existing Work

Super-resolution techniques may be used to perform image-level fusion. Super-resolution is the process of generating an image with a higher resolution than the corresponding source images. The information from individual frames can be fused into a single composite image with higher resolution, resulting in better recognition performance. Although much work has been done in the face recognition domain, super-resolution of iris images has not been widely discussed in the literature. This is due to the stochastic nature of the iris texture which does not lend itself to traditional super-resolution schemes. Fahmy [99] describes an interleaving process to generate a high resolution iris image from a low resolution face video. Iris regions of equal sizes are segmented from the low-resolution frames of a face video. These iris regions are registered using a cross correlation model, and interleaved with each other to form an image of higher resolution. This process is iterated multiple times to generate a high resolution iris image.

Huang et al. [100] propose a learning based algorithm to improve the resolution of normalized iris images. The algorithm is trained using a large number of image pairs consisting of low-resolution normalized iris images and their corresponding high-resolution versions. In the training stage, each low-resolution normalized iris image is tessellated into multiple blocks, and the relation of each block with its corresponding high-resolution pair is modeled using Markov networks. In the testing stage, a high-resolution output is generated by upsampling the input low-resolution image and restoring the lost frequency information based on the best matching training blocks from the database.

2.2.2 Fusion in Iris Videos

While the input in the above mentioned approaches is a static set of individual images, Hollingsworth et al. [101] use a set of frames extracted from an iris video. A set of 10 good quality frames are chosen automatically and the iris is segmented and normalized in

each of them. These normalized irides are then fused on a pixel-by-pixel basis, by using an operator (e.g., *mean*, *median*, etc.). Consider a set of n images $\{I_1, I_2, \dots, I_n\}$ each of size $M \times N$. The intensity of the final fused image I_{fused} at a location (i, j) can be obtained via the mean operator as

$$I_{fused}(i, j) = \frac{1}{n} \sum_{p=1}^n I_p(i, j), \quad (2.1)$$

where $1 \leq i \leq M$ and $1 \leq j \leq N$. This technique can be viewed as a *pixel-level* fusion scheme, where the pixel intensity at a given location of the output is dependent only on the corresponding pixel intensities of the input images. Although the technique is simple, the recognition performance of the resulting output is greatly improved [101]. However, such an output strongly depends on the following factors:

1. *Number of observations*: The output is typically more reliable if the number of the input samples, n , is large. If n is small, the output can be a weak estimate.
2. *Accuracy of observations*: If a majority of the input images contain noise, the quality of the output image cannot be expected to improve over the input images. For example, if a large number of input images are blurred at a specific region, it cannot be rectified in the resulting output.

Furthermore, the input images should be perfectly registered. If the registration is inaccurate, the output would be an approximate or a smoothed representation of the actual scene. Perfect registration of iris images obtained in non-ideal environments is a challenging problem. In iris images, imperfect registration can perturb the texture of the output and reduce the matching performance of the system. Thus, it has to be ensured that the fusion scheme does not alter the textural richness of the iris.

The approach proposed in this thesis performs image-level fusion in two stages. In the first stage, an image reconstruction scheme based on the Principal Components Transform is used to re-estimate the input images. In the second stage, the reconstructed input images are fused by an image averaging scheme using the *mean* operator. The recognition performance obtained using the output generated by the two-stage approach is observed to be higher than the performances obtained after applying either of the stages.

2.3 Principal Components Transform

2.3.1 PCT versus PCA

Principal Components Analysis (PCA) has been widely used in the field of automatic face recognition. Turk and Pentland [102] view each face image as a point in a high-dimensional sub-space whose coordinates correspond to the pixel intensities. Based on a large set of registered training face images, PCA is used to determine a set of orthogonal basis vectors, referred to as *eigenfaces* or *principal components*, that correspond to directions of maximum variance in the original sub-space. Subsequently, any given face image can be represented as a weighted sum of such eigenfaces. In the field of iris recognition, Dorairaj et. al. [82] use PCA to determine a set of basis vectors for iris images. However, in the absence of a common morphology in normalized iris (unlike face which has some common landmarks across images), the basis vectors do not have a trivial physical interpretation.

In this thesis, PCA is used in a different manner than what has been typically used in the biometrics literature. To avoid confusion, the PCA technique used in [102] (and other publications) is referred to as the *conventional PCA*, while the technique used in this work is referred to as *Principal Components Transform (PCT)*. The major differences between the conventional PCA, and the PCT approach used in this work are listed below:

1. Given a set of n images, each having a spatial resolution of $M \times N$, conventional PCA represents every image as a point in the MN dimensional space. In this work, an n dimensional space is considered in which each *pixel intensity vector* is a point. The pixel intensity vector, \vec{V}_{pq} , that contains pixel intensity values across all given images at a location (p, q) , is defined as:

$$\vec{V}_{pq} = [I_1(p, q), I_2(p, q), \dots, I_n(p, q)]^T, \quad (2.2)$$

where $I_j(p, q)$ denotes the pixel intensity value of an image I_j at a location (p, q) , and $j = \{1, \dots, n\}$, $p = \{1, \dots, M\}$, $q = \{1, \dots, N\}$. This variation in representation can be easily obtained by considering a different scheme for stacking (or arranging) the pixel values of the images, as shown in Figure 2.2.

2. Conventional PCA is typically applied on multiple images of different subjects, to highlight the variance information among the images. On the other hand, PCT is

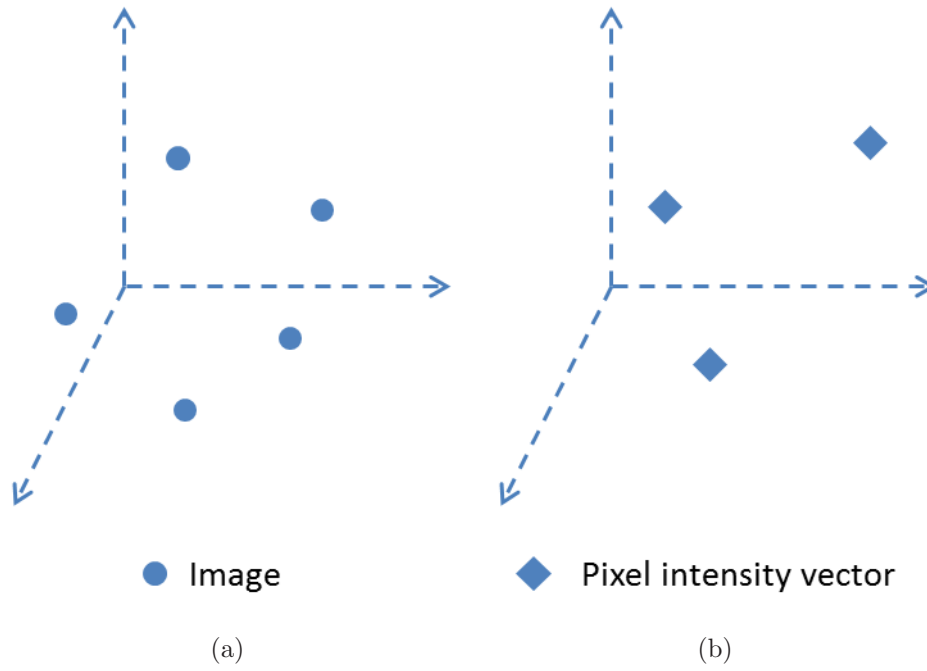


Figure 2.2: Difference between PCA and PCT: (a) In an MN dimensional space, the conventional PCA approach considers each image as a point (denoted by a circle). (b) In an n dimensional space, the PCT approach considers every pixel intensity vector as a point (denoted by a diamond).

applied on multiple images of the same subject, to highlight the variance information among the pixel intensity vectors. Consequently, PCT seeks to extract *discriminatory* pixels from the iris frames.

2.3.2 Mathematical Formulation

Consider a set of n images $\{I_1, I_2, \dots, I_n\}$, each having a spatial resolution of $M \times N$ pixels. Every image I_j , is transformed to a row vector, \vec{I}_j , of size $1 \times MN$ where $j = \{1, 2, \dots, n\}$. An image data matrix \mathbf{X} is obtained by stacking* the n row vectors, one per row, as shown below:

$$\mathbf{X} = \begin{bmatrix} \vec{I}_1 \\ \vec{I}_2 \\ \vdots \\ \vec{I}_n \end{bmatrix}. \quad (2.3)$$

*The major difference between the conventional PCA and the PCT approach lies in the stacking process used to generate the image data matrix. Turk and Pentland [102] stack the images as column vectors into \mathbf{X} , whereas PCT considers the images as row vectors. See Figure 2.3 for a better understanding of the stacking process.

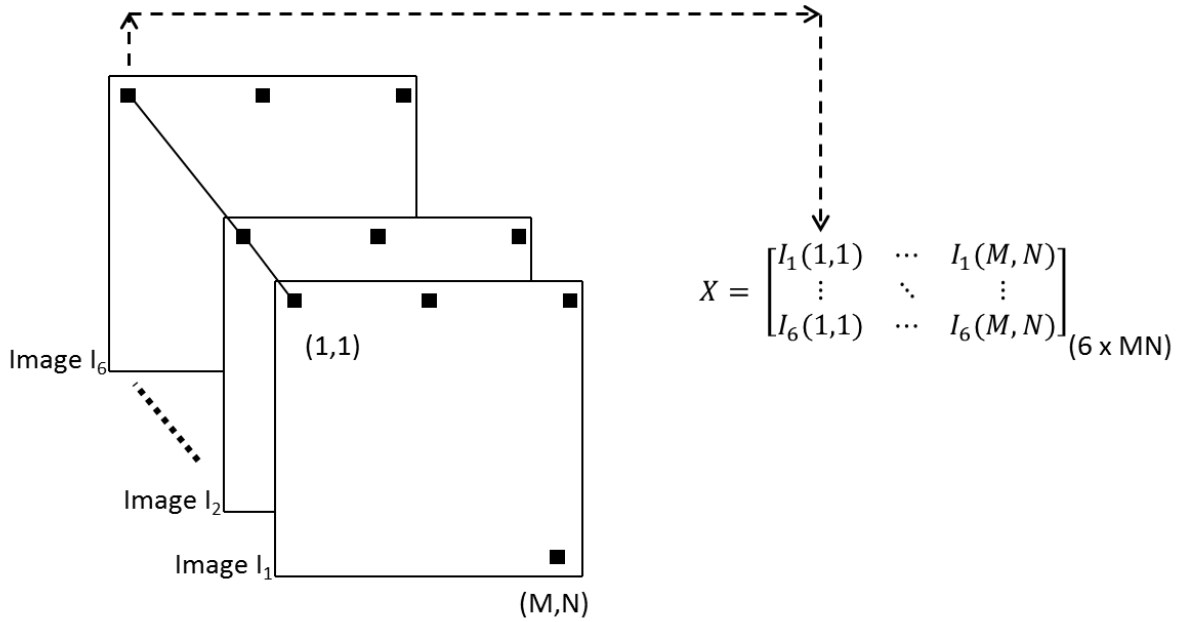


Figure 2.3: The process of stacking image pixels used to generate image data matrix \mathbf{X} for the PCT approach.

The size of \mathbf{X} will be $n \times MN$. In other words, every column of the image data matrix \mathbf{X} is a *pixel intensity vector* at a particular location. This process is illustrated in Figure 2.3. For this data, the empirical mean vector $\vec{\mathbf{m}}_{\mathbf{X}}$ is computed along each dimension. The resulting row vector of size $1 \times MN$ is given by the following equation:

$$\vec{\mathbf{m}}_{\mathbf{X}} = \frac{\sum_{p=1}^n \vec{I}_p}{n}. \quad (2.4)$$

The covariance matrix $\mathbf{C}_{\mathbf{X}}$ for the image data can be computed by the equation:

$$\mathbf{C}_{\mathbf{X}} = \frac{1}{n}(\mathbf{X} - \mathbf{I}\vec{\mathbf{m}}_{\mathbf{X}})(\mathbf{X} - \mathbf{I}\vec{\mathbf{m}}_{\mathbf{X}})^T \quad (2.5)$$

where \mathbf{I} is an identity matrix of size $n \times 1$. The size of the covariance matrix $\mathbf{C}_{\mathbf{X}}$ will be $n \times n$.

The eigenvectors of the covariance matrix $\mathbf{C}_{\mathbf{X}}$ are obtained by decomposing it into its canonical form. Using this information, the input data can be transformed into a new feature space by the equation:

$$\mathbf{Y} = \mathbf{A}(\mathbf{X} - \mathbf{I}\vec{\mathbf{m}}_{\mathbf{X}}), \quad (2.6)$$

where \mathbf{A} is an $n \times n$ matrix, whose rows contain the normalized eigenvectors of $\mathbf{C}_{\mathbf{X}}$. At any point of time, the original data \mathbf{X} can be recovered by performing a simple inverse

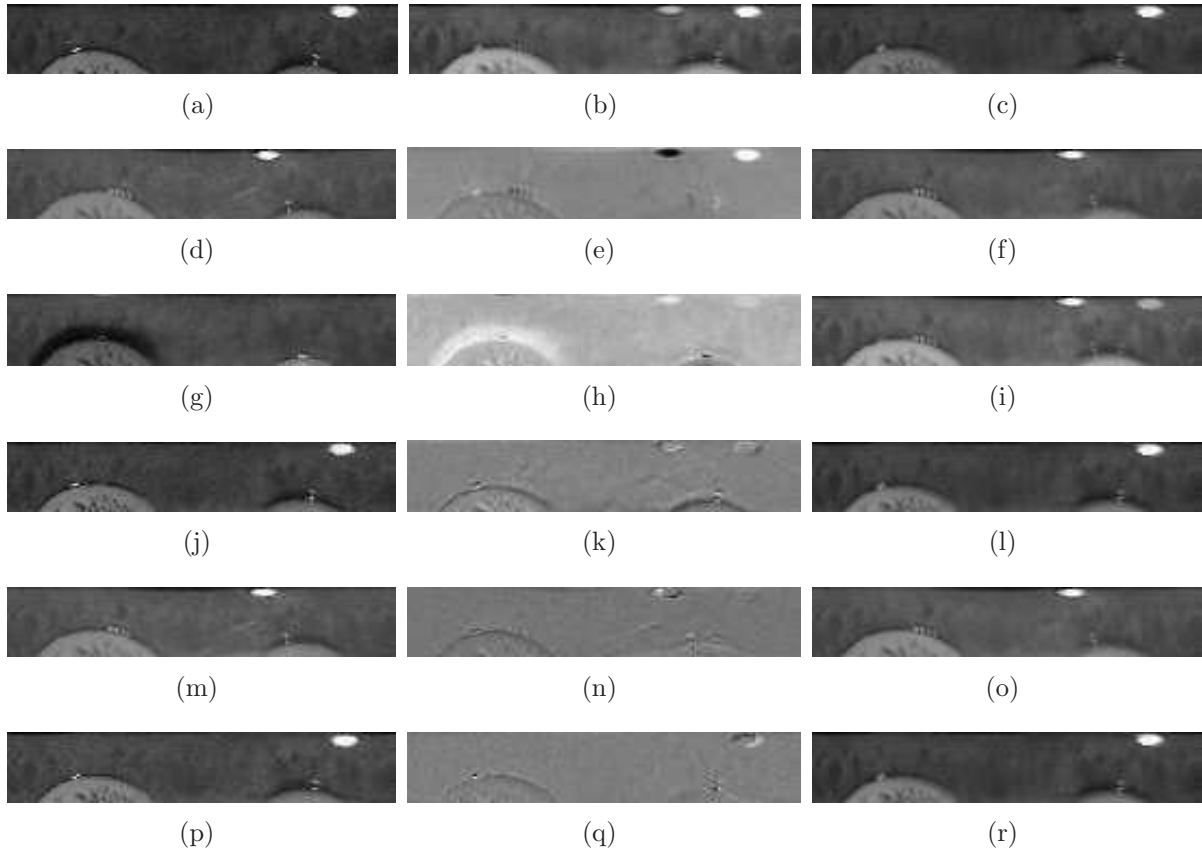


Figure 2.4: Left column: Normalized input iris images. Middle column: principal components of the input data arranged in the decreasing order of magnitude. Right column: reconstructed output obtained by using the PCT approach with $q = 2$. Images have been scaled to fit the document.

transformation given by:

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{Y} + I\vec{m}_{\mathbf{X}}. \quad (2.7)$$

As the rows of \mathbf{A} are ortho-normal vectors, $\mathbf{A}^{-1} = \mathbf{A}^T$. Hence the above equation becomes:

$$\mathbf{X} = \mathbf{A}^T\mathbf{Y} + I\vec{m}_{\mathbf{X}}. \quad (2.8)$$

If only the most significant q principal components of the data are retained, \mathbf{A} becomes a $q \times n$ matrix, denoted as \mathbf{A}_q . The transformed data obtained by using only the selected q principal components is given by the following equation:

$$\tilde{\mathbf{Y}} = \mathbf{A}_q(\mathbf{X} - I\vec{m}_{\mathbf{X}}), \quad (2.9)$$

where the size of $\tilde{\mathbf{Y}}$ is $q \times MN$. Even though some of the eigenvectors are discarded, it is still possible to recover the initial input data by considering the following approximation:

$$\tilde{\mathbf{X}} = \mathbf{A}_q^T\tilde{\mathbf{Y}} + I\vec{m}_{\mathbf{X}}, \quad (2.10)$$

Consider an input of 6 normalized iris images, as shown in the left panel of Figure 2.4. The principal components of this data, arranged in the descending order of magnitude, are shown in the central panel of Figure 2.4. It can be noticed that the principal components of higher magnitude account for the maximum variability in the input data. If all the principal components are used, the original data can be reconstructed. The images reconstructed using only the top q (in this case, $q = 2$) principal components are shown in the right panel of Figure 2.4. The above process results in an approximation of the initial data, since all the eigenvectors are not used during reconstruction. However, the impact of dropping the lowest-valued eigenvectors is less significant on the reconstructed data.

2.4 Proposed Approach

The following steps describe the process by which the proposed image-level information fusion scheme is applied to low resolution iris video frames:

1. Let V be a low-resolution iris video containing n frames denoted by $F = \{f_1, f_2, \dots, f_n\}$.
2. A set of k good quality frames are manually selected from the available frame set. It is not necessary for the selected frames to be successive in the video stream.
3. The selected frames are processed to segment and normalize (un-wrap) the iris to equal sized rectangular entities.
4. PCT is applied on these normalized frames to obtain the reconstructed frames [†].
5. The reconstructed frames, are further fused by image averaging process, yielding a single output image. This output, instead of the original input, is used during the recognition process.

By using the evidence of multiple frames, the PCT based on pixel intensity vectors, projects the iris image onto a lower manifold where its discriminatory information is optimized. This optimization is accomplished using the eigenvectors of the covariance matrix of pixel intensities. While other types of manifold analysis techniques can be used, in the

[†]The transformation alters the content of the input frames but does not change their spatial resolution.

interest of computational complexity, the PCT scheme is adopted in this work. By reconstructing the frames based on the eigenvectors, the salient information is extracted. A schematic of the technique is shown in Figure 3.6.

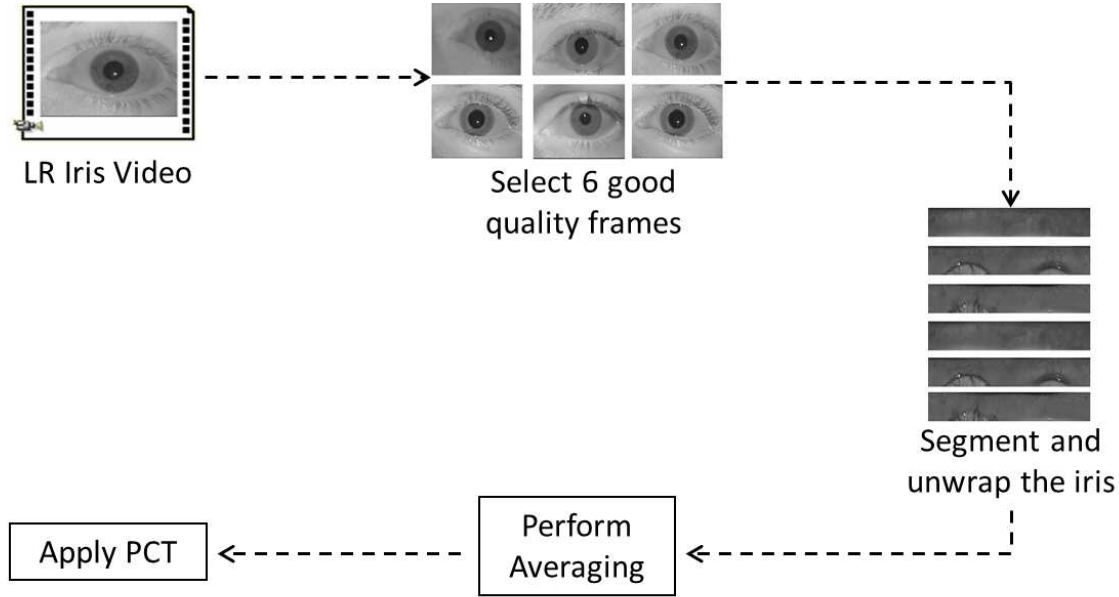


Figure 2.5: Proposed image-level fusion scheme.

2.5 Experiments and Results

2.5.1 Database

A subset of the Multi-Biometric Grand Challenge (MBGC) database [103] containing Near Infrared (NIR) iris videos was used for the experiments. The iris video streams of multiple subjects are recorded in MPEG-4 format under varying illumination. A set of 110 right iris videos were selected for the study by considering 1 video each of 110 unique subjects. Frames are extracted from the videos and saved in BMP format without any compression. Each video contained 300 frames on an average, with every frame having a spatial resolution of 640×480 pixels.

A gallery set comprising a total of 440 images was formed by selecting 4 frames per subject. The value of k , which represents the number of low-resolution probe frames, was chosen as 6. This provides for a direct comparison of the recognition performances obtained by the proposed approach and the image averaging approach. [101]. Therefore, a set of 660

frames (6 frames each, for 110 videos) were chosen as the probe set. The frame selection process was performed manually, based on factors that impact the amount of information useful for recognition (e.g., percentage of iris visible in the frame, specular reflection, blur, occlusion due to eyelashes and eyelids, etc.).

2.5.2 Pre-processing

To *simulate* the low-resolution imagery, the original probe set of 640×480 pixel resolution (referred to as *ProbeSet L1*) is sub-sampled. Sub-sampling was performed by an averaging operator to reduce the size of a frame by a factor of 4: $1/2$ the length and $1/2$ the width. The sub-sampling operation was performed on every frame of the *ProbeSet L1* to obtain a lower-resolution frameset. This process was used iteratively to generate multiple framesets of the following resolutions: 320×240 , 160×120 , and 80×60 (referred to as *ProbeSet L2*, *ProbeSet L3*, and *ProbeSet L4*, respectively). Any resolution below 80×60 pixels (*ProbeSet L4*) was considered too low to work with. The average diameter of the iris in *ProbeSet 1*, *ProbeSet 2*, *ProbeSet 3*, and *ProbeSet 4* was approximately 220, 110, 50, and 20 pixels, respectively. Figure 2.6 shows a sample frame at various resolutions.

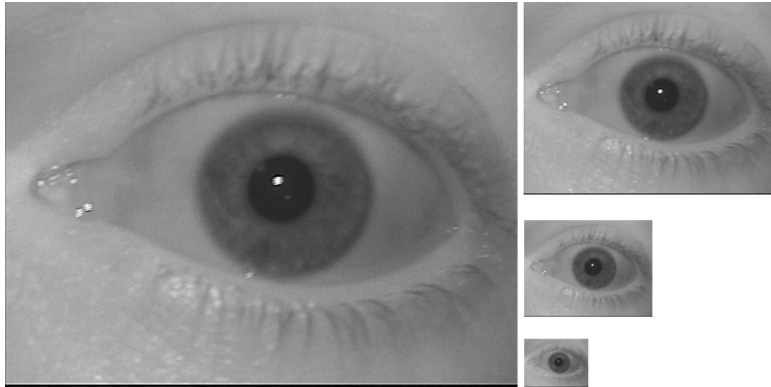


Figure 2.6: A sample right iris frame of resolution 640×480 down-sampled to resolutions of 320×240 , 160×120 , and 80×60 using the image averaging operator. Images have been scaled to fit the document.

2.5.3 Iris Segmentation

Three different approaches were used to perform automatic iris segmentation[‡]: (a) Hough transform, (b) Integro-differential operator, and (c) Geodesic Active Contours.

[‡]A detailed description of the three considered approaches is provided in Appendix A.

Segmentation technique	Number of input images	Number of correctly segmented images	Segmentation accuracy
Integro-differential operator	660	353	53.4%
Hough transform	660	371	56.2%
Geodesic Active Contours (GAC)	660	427	64.6%

Table 2.1: **Segmentation accuracies of the three techniques using ProbeSet L4.**

The segmentation accuracies[§] of the three approaches using *ProbeSet L4* are listed in Table 2.5.3. It can be observed that the segmentation performance of all the considered approaches is low. Segmenting the iris becomes increasingly difficult with the decreasing resolution of the image. This is due to the reduced quality of a low-resolution image, thereby impacting the iris boundary information.

To avoid the effect of incorrect iris segmentation on the recognition performance, a semi-automated segmentation scheme is adopted. Iris segmentation was performed independently on each frame for the various probe sets. This is performed by a human observer by marking the boundaries of the iris in a frame. Simultaneously, a noise mask is created, which records the locations of eyelids and eyelashes that occlude the segmented iris. Both the segmented iris region and the noise masks are normalized using Daugman’s rubber sheet model [62]. For the experiments, the segmented irides in all probe sets (varying resolution) are normalized to a fixed size: 32×180 pixels. Usually, the most suitable size for normalizing an iris image is based on the radii of the pupil and iris. However, matching normalized irides of different sizes is not possible using the existing Daugman’s approach [62]. In general, matching iris images of different resolutions is still an open problem. The loss of textural information caused by normalizing frames of varying resolutions to a fixed size can be noticed in Figure 2.7.

To extract the textural features of the iris, a two dimensional Gabor filter is convolved with the unwrapped iris image. The output of this convolution operation contains both the real and imaginary responses. A phase demodulation process is used to encode these responses to a binary biometric template, often called as an *IrisCode*. Hamming distance

[§]Segmentation accuracy = $\frac{\text{Number of correctly segmented images}}{\text{Number of input images}} \times 100$

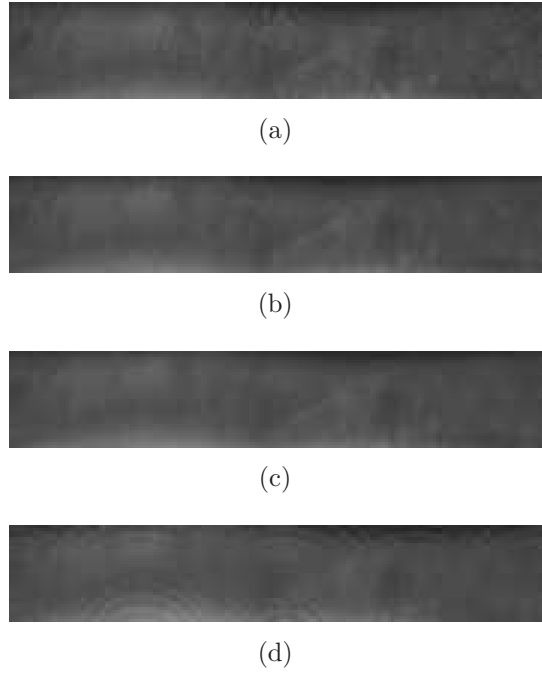


Figure 2.7: Normalized probe images corresponding to an iris frame at multiple resolutions: (a) 640×480 , (b) 320×240 , (c) 160×120 , and (d) 80×60 . Images have been scaled to fit the document.

is used to measure the dissimilarity between two *IrisCodes* while masking the corresponding noisy regions. An open source MATLAB implementation [77] was used with minor modifications to perform above mentioned operations.

2.5.4 Recognition Accuracy

Two iris recognition software packages, IrisBEE [104] and VeriEye [105], were initially used to observe the matching performance at various image resolutions. However, neither package could generate the iris templates or perform matching on probe images below a resolution of 320×240 . Thus, the performance evaluation in this work is conducted using an open source MATLAB implementation [77] for iris encoding and matching. Receiver Operating Characteristic (ROC) curves are used to evaluate and compare recognition performance. Every frame in the probe set is matched against all the gallery frames. A total of 2,640 genuine scores and 287,760 impostor scores were obtained by the matching process for each probe set. The performances obtained by matching the probe sets with the gallery are shown in Figure 2.8.

From Figure 2.8, it is observed that the recognition performance drops significantly

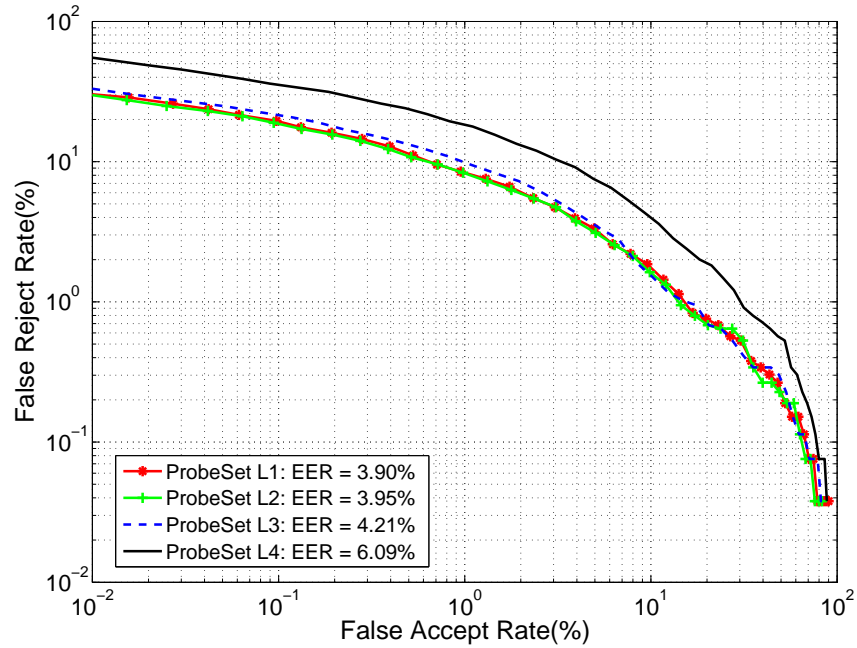


Figure 2.8: ROC curves for the right iris data, obtained by matching probe sets with the gallery.

for *ProbeSet L4*. Therefore, the proposed algorithm is used on this set to improve its performance. As interpolation is a commonly used technique to upsample low-resolution images, the matching performance obtained after interpolation is used as the baseline. Hence, the recognition performance obtained by interpolating *ProbeSet L4* to a resolution of 160×120 is used as a baseline. Figure 2.9 shows the recognition performances of the new framesets obtained by applying (a) only PCT, (b) only image averaging, and (c) the proposed approach, on normalized frames of *ProbeSet L4*. From the figure, it can be noticed that the Equal Error Rate of *ProbeSet L4* is reduced from 6.09% to 1.76%. This suggests that the recognition performance of the proposed approach is better than performances obtained by the individual stages involved (PCT and averaging).

The genuine and impostor match score distributions of *ProbeSet L4* before and after applying the proposed technique are shown in Figure 2.10. From the results, it is observed that the genuine match score distribution shifts toward zero, indicating a reduction in false reject rate (FRR).

To further evaluate the performance of the proposed technique, the above experiment was repeated with the same setup using the left iris videos of 100 subjects. The perfor-

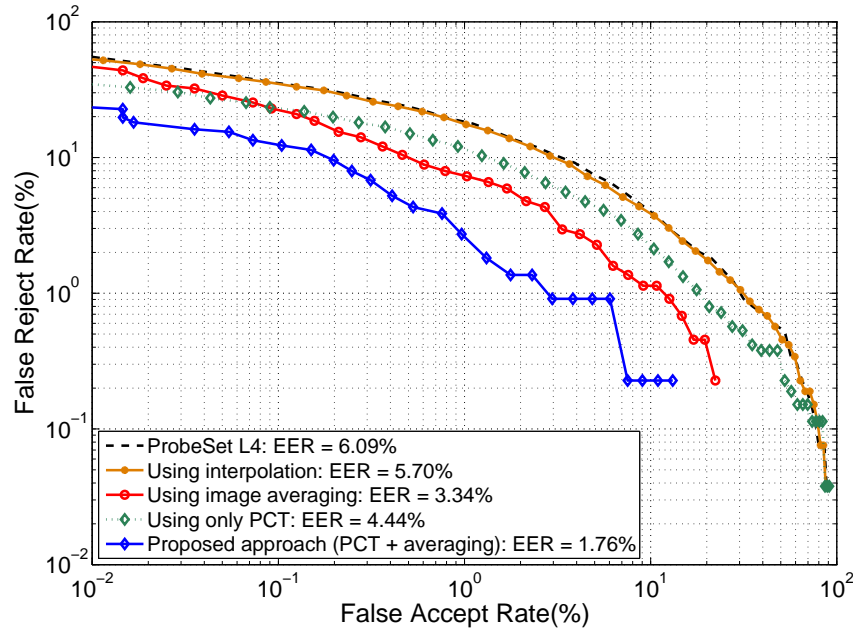
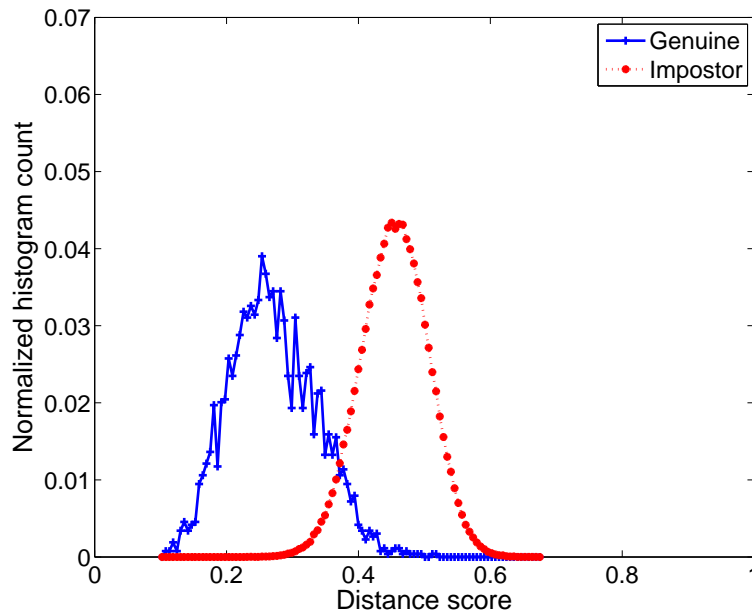


Figure 2.9: ROC curves for the right iris data, obtained before and after applying the proposed technique. Note that by applying the proposed approach, the Equal Error Rate of *ProbeSet L4* is reduced from 6.09% to 1.76%.

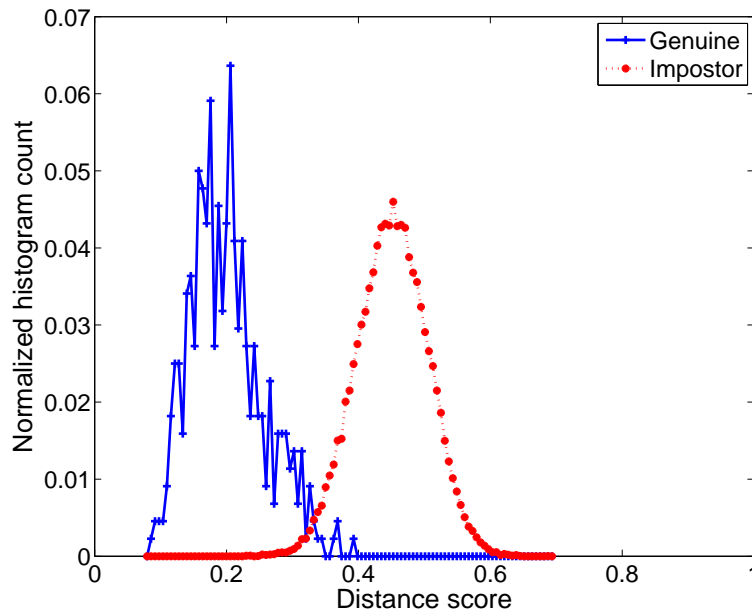
mances obtained using this data are summarized in Table 2.2. From the results obtained using both left and right iris videos, it can be stated that the recognition performance of low resolution frames can be significantly improved using the proposed technique. Additionally, it is observed from Table 2.2 that the recognition performance is slightly improved when images are down-sampled from a resolution of 640×480 to 320×240 . One of the reasons for such effect could be the reduction of noise by the down-sampling process. Similar observations were made in studies related to the impact of fingerprint resolution on recognition performance [106].

Table 2.2: Equal Error Rates obtained using the left iris video data.

Probe Set	EER
ProbeSet L1	3.16%
ProbeSet L2	2.96%
ProbeSet L3	3.10%
ProbeSet L4	5.45%
On interpolating ProbeSet L4	3.38%
On applying PCT on ProbeSet L4	2.03%
On applying averaging on ProbeSet L4	2.58%
On applying the proposed approach (PCT + averaging) on ProbeSet L4	1.48%



(a)



(b)

Figure 2.10: Genuine and impostor match score distributions for the right iris data: (a) before, and (b) after applying the proposed technique.

2.5.5 Proposed Approach on High-Resolution Images

From Figures 2.8, 2.9, and Table 2.2, it can be observed that the performance obtained by applying the proposed approach on *ProbeSet L4*, is better than that of *ProbeSet L1*.

This observation can lead to an assumption that fusing the low-resolution images by the proposed approach can provide better performance than the high-resolution images. However, it has to be noted that the significant improvement in the performance of *ProbeSet L4* is caused mainly by the image averaging process. This is because the image averaging step yields a significantly lower number of images for *ProbeSet L4* upon applying the proposed approach (110 and 100 images for right and left sides, respectively), when compared to the number of images in *ProbeSet L1* (660 and 600 images for right and left sides, respectively). To validate this argument, the proposed approach (PCT + averaging) is also applied on *ProbeSet L1*. The Equal Error Rates obtained using left and right iris *ProbeSet L1* data are listed in Table 2.3. The corresponding ROC curves for right iris data are shown in Figure 2.11. From the results, it can be observed that the proposed approach greatly improves the performance of *ProbeSet L1*, in comparison to that of *ProbeSet L4*. The contribution of image averaging in improving the recognition performance can be clearly observed from Table 2.3. These results suggest that the proposed approach can improve the performance of any given set of iris video frames.

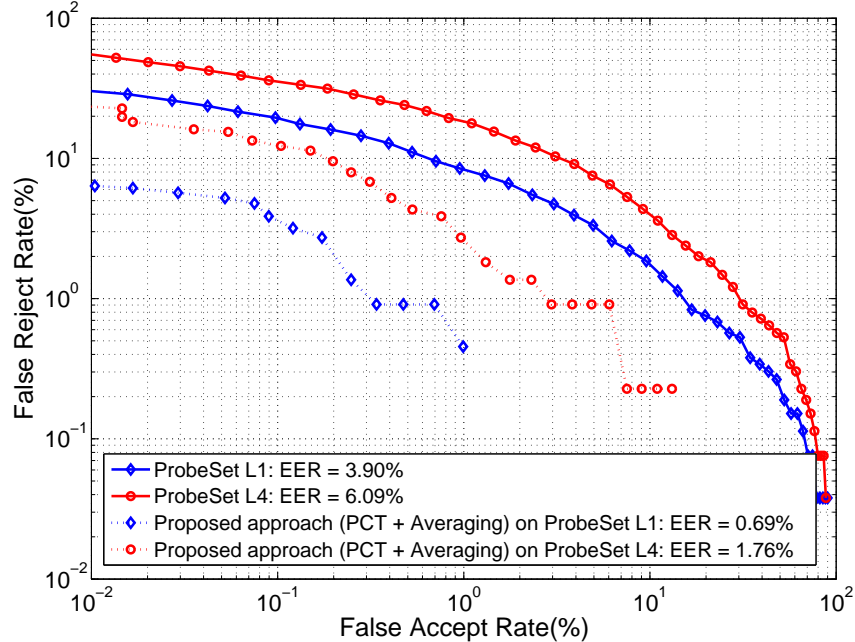


Figure 2.11: ROC curves for right iris *ProbeSet L1* and *ProbeSet L4*, before and after applying the proposed approach.

Table 2.3: **Equal Error Rates obtained before and after applying the proposed approach on right and left iris *ProbeSet L1* data.**

Probe Set	Right Iris	Left Iris
ProbeSet L1	3.90%	3.16%
On applying image averaging on ProbeSet L1	1.30%	1.56%
On applying PCT on ProbeSet L1	1.75%	1.89%
On applying the proposed approach (PCT + averaging) on ProbeSet L1	0.69%	0.82%

2.5.6 Effect of Down-sampling Methods

As described in Section 2.5.2, the low-resolution probe frames used in this work were generated by an image averaging operator. The averaging operator sub-samples a given frame by a fixed factor of 4: $1/2$ the length and $1/2$ the width. A potential concern with this approach is whether the output obtained by such a process accurately represents low-resolution imagery encountered in practical scenarios (e.g., when using low-resolution optics or sensors, or when large stand-off distances are involved). The most accurate approach to address this concern would be by using the exact transfer function (or, down-sampling method) that relates the high- and low-resolution images. However, the process of learning the actual transfer function is very complicated, and an area of research by itself. Furthermore, sensor parameters for biometric images are not always known (e.g., when using surveillance data). In this regard, a Gaussian filtering approach was used to down-sample the images. The Gaussian filter, due to its anti-aliasing property, is considered to provide an output that is closer to real-world low-resolution imagery [107]. Gaussian filters are the only class of functions whose point spread function (PSF) and optical transfer function (OTF) belong to the same class. This is because the Fourier Transform of a Gaussian function is still a Gaussian.

The experiments described in Section 2.5.4 were repeated by *simulating* the low-resolution images using a Gaussian filter. The images obtained by Gaussian down-sampling are shown in Figure 2.12. Note that the images are scaled to fit the document and it is difficult to assess the variations in image quality by comparing Figure 2.12 with Figure 2.6. To provide a visual comparison of the two down-sampling techniques, normalized images obtained by Gaussian and averaging filters are shown in Figure 2.13. From the figure, it can be observed

that the visual variations between the images are very minimal. The recognition performances obtained using images down-sampled by Gaussian filtering are listed in Table 2.4. From the results, it can be observed that the Equal Error Rates obtained using images down-sampled by the Gaussian approach are not significantly different from those obtained using averaging filter. In most cases, the EERs corresponding to Gaussian down-sampled images are slightly higher. This is because the anti-aliasing property of the Gaussian induces blur in the images, which impacts the recognition performance. From the results, it can be observed that the proposed approach improves the recognition performance even in the considered case.

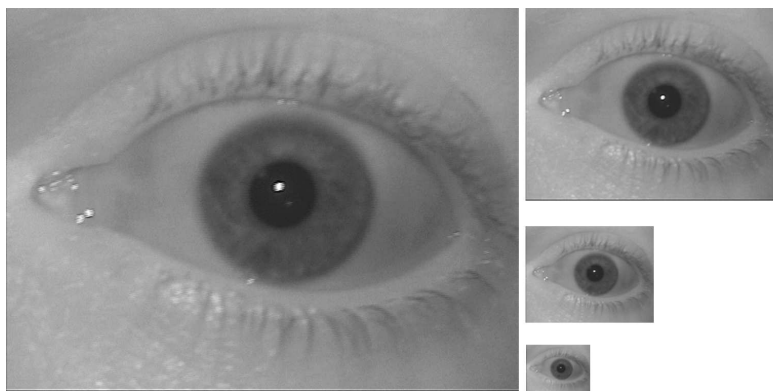


Figure 2.12: A sample right iris frame of resolution 640×480 down-sampled to resolutions of 320×240 , 160×120 , and 80×60 using the Gaussian filtering approach. Images have been scaled to fit the document. Note that the same image was used in Figure 2.6.

Table 2.4: **Equal Error Rates obtained before and after applying the proposed technique on right and left iris *ProbeSet L1* data downsampled by the Gaussian operator.**

Probe Set	Right Iris	Left Iris
ProbeSet L1	3.90%	3.16%
ProbeSet L2	3.99%	3.00%
ProbeSet L3	4.32%	2.38%
ProbeSet L4	6.24%	5.38%
On applying the proposed approach (PCT + averaging) on ProbeSet L4	1.80%	1.56%

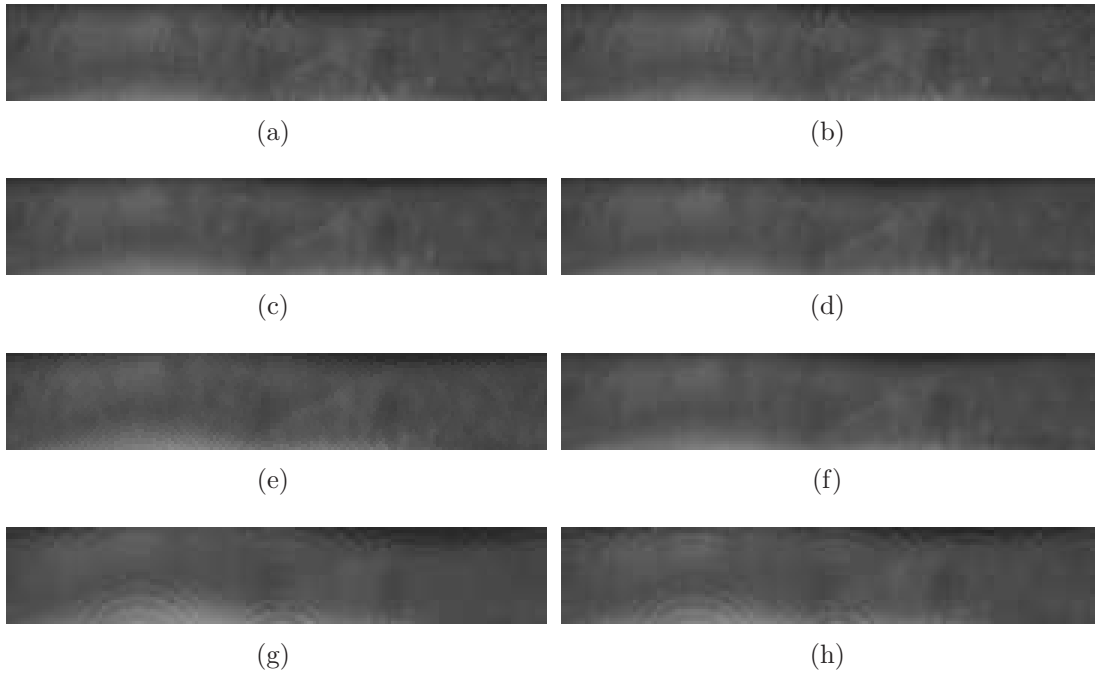


Figure 2.13: Normalized images obtained from *ProbeSet L1* to *L4* that are downsampled using Gaussian ((a), (c), (e), and (g)), and averaging ((b), (d), (f), and (h)) operators. Images have been scaled to fit the document.

2.6 Score-Level Fusion

The match scores generated by comparing a gallery image against a multi-frame probe set can be fused by employing a score-level fusion scheme [98]. In the current framework, score-level fusion is applied to (a) *ProbeSet L4* and (b) the corresponding frame set obtained after applying PCT. Given a set of match scores $\{S_1, S_2, \dots, S_n\}$ obtained by matching n probe frames $\{f_1, f_2, \dots, f_n\}$ against a gallery image, a new score is generated by the sum rule that merely takes the average of these scores. The EERs obtained by invoking score-level fusion on the right and left iris video data are summarized in Table 2.5. The score-level fusion of the PCT output reduces the Equal Error Rate of *ProbeSet L4* from 6.09% to 1.45% for the right iris videos, and from 5.45% to 1.46% for the left iris videos, respectively. The results indicate that the recognition performance can be further improved by using score-level fusion.

Table 2.5: **Equal Error Rates obtained before and after applying score-level fusion on the right and left iris video data.**

	Right Iris (110 subjects)		Left Iris (100 subjects)	
	Before	After	Before	After
ProbeSet L4	6.09%	2.59%	5.45%	2.25%
PCT frameset	4.44%	1.45%	2.03%	1.46%

2.7 Effect of Image Registration

The image frames used in this work were extracted from continuous iris videos. All the videos were acquired in a controlled environment with minimal eye movements and illumination variation. As a result, the need for a computationally expensive image registration step can be reduced. Under unconstrained conditions, however, it cannot be expected that the input images are well registered. Occlusions, eye movements, and contraction or dilation of the pupil caused by illumination variations, can introduce significant image registration errors. A robust image level fusion scheme should successfully handle such variations to provide a reliable recognition performance. In this regard, the effectiveness of the proposed approach is tested using misaligned images. To this end, each of the 6 unwrapped iris images corresponding to every subject within *ProbeSet L4* is subjected to a random shift[¶] of: (a) 10 pixels and (b) 15 pixels. Sample images containing such shifts are shown in Figure 2.14. The choice of image shifts of 10 or more pixels is based on the fact that Libor Masek’s implementation of Daugman’s algorithm compensates for image shifts of up to 8 pixels. The corresponding noise masks are also shifted accordingly, and recognition performance is computed.

The EERs obtained after shifting the individual frames, and after applying the proposed approach are listed in Table 2.6. The EERs obtained after applying the image averaging approach [101] are also provided for reference. From the results, it can be observed that both the techniques provide poor performance when images are misaligned. This indicates the need for robust image registration before applying both the approaches.

[¶]A random shift refers to any one of the following image shifts: up, down, left, and right.

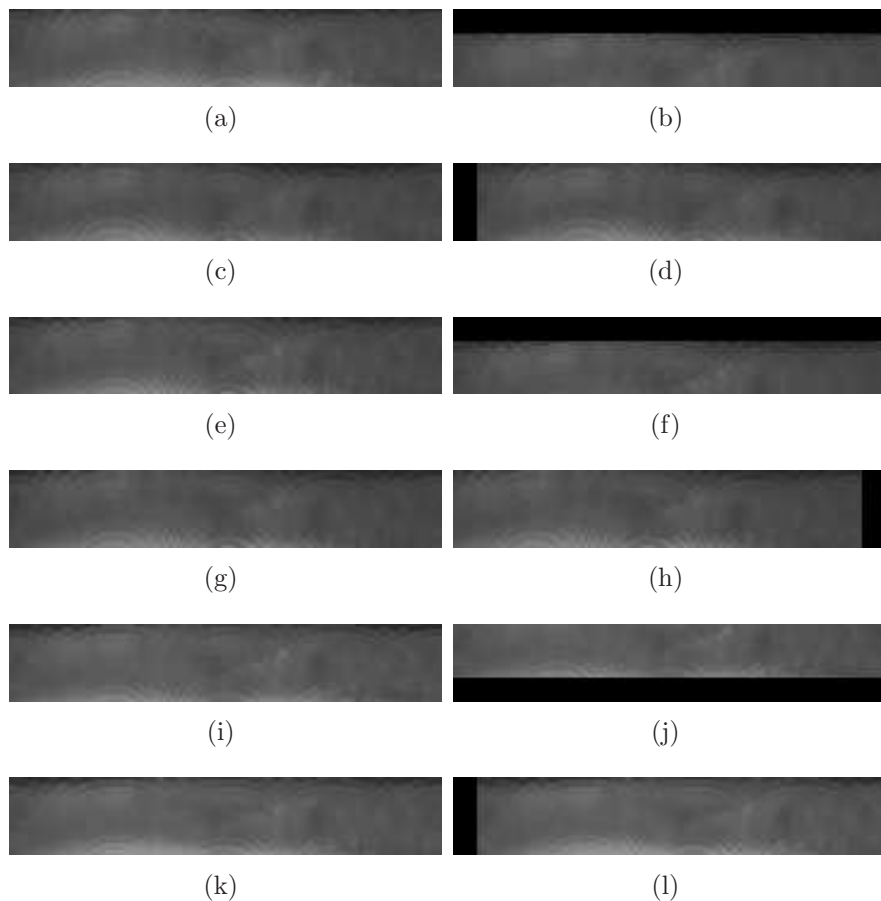


Figure 2.14: Normalized images corresponding to a single subject from *ProbeSet L4* before (left column) and after (right column) applying a random 10 pixel shift. Images have been scaled to fit the document.

Table 2.6: Equal Error Rates obtained after image shifting and after applying the proposed technique on right iris *ProbeSet L4* data.

Shift	After applying shift	After image averaging	After applying the proposed approach
10 pixels	11.0%	32.2%	31.0%
15 pixels	23.2%	36.2%	36.7%

2.8 Summary

An image-level fusion scheme is proposed which improves the recognition performance of low-resolution iris images. By using the proposed approach on a low-resolution iris database (average iris diameter of 20 pixels), the equal error rates are significantly reduced (from 6.09% to 1.76% for the right iris and from 5.45% to 1.48% for the left iris, respectively). The effect of using two different down-sampling methods is also investigated, and the proposed approach provides an improved performance in both considerations. The use

of complex routines to enhance low-resolution iris videos can be avoided using the proposed approach. Possible extension of this work would include further improving the recognition performance by considering other sub-space analysis techniques for fusion (such as Linear Discriminant Analysis, Independent Component Analysis, etc.). The performance of the proposed technique when the input images are not properly registered could also be investigated. A large scale evaluation on larger and more challenging databases can be performed.

Chapter 3

Plastic Surgeries: Face & Ocular Biometric Fusion

3.1 Introduction

The field of automatic face recognition has been significantly researched over the past 20 years. Extensive research efforts in this area have helped in attaining high recognition performances [108]. However, the problem of automatic face recognition is not completely solved yet. This is because the face recognition performance is negatively impacted in the presence of occlusions and pose, illumination, expression (PIE) variations [109]. More recently, a broader range of problems have garnered the interest of face recognition researchers. Some of these problems include: photo-to-sketch matching, caricature recognition, age invariant recognition, matching plastic surgery images, etc. [110].

The task of successfully matching face images obtained before and after plastic surgery is a challenging problem. The degree to which a face is altered depends on the type and number of plastic surgeries performed, and it is difficult to model such variations. In this chapter, a fusion approach is proposed that combines information from the face and ocular regions to enhance recognition performance in the identification mode. It is observed that the proposed approach provides the highest reported recognition performance (at the time this work was done) on a publicly accessible plastic surgery database. Compared to existing approaches, the proposed approach is not learning based and this reduces computational requirements. This chapter also presents a systematic study of the matching accuracies corresponding to various types of surgeries.

3.2 Facial Plastic Surgeries

Facial plastic surgery generally refers to a medical procedure that involves modifying the appearance of external facial anatomical features using surgical methods [111]. Based on their purpose, plastic surgeries can be broadly classified into two categories:

1. **Reconstructive:** These surgeries are performed mainly to reconstruct the generic appearance of a facial feature, so that its functionality is restored or improved. For example, surgical treatment of ptosis (drooping of the upper eyelid due to weak muscles, that can cause vision interference).
2. **Aesthetic improvement:** These surgeries are performed to alter the appearance of a fully functional feature, solely with the purpose of aesthetic improvement. For example, restoring damaged skin due to burn injuries or accidents.

Facial plastic surgeries have become increasingly popular in the recent past, especially for aesthetic improvement purposes. A report from the American Society of Plastic Surgery states that a total of 14.6 million cosmetic and reconstructive plastic surgeries were performed just within USA in the year 2012 [112]. Three of the top five surgeries in this set relate to the modification of facial features [113]. Some of the major facial plastic surgeries include: rhinoplasty (nose surgery), blepharoplasty (eyelid surgery), brow lift (eyebrow surgery), otoplasty (ear surgery), and rhytidectomy (face lift surgery) (see Figure 3.1). A detailed, but non-exhaustive list of facial plastic surgeries is provided in [114].

The degree to which the appearance of a human face can be modified by plastic surgery, depends on the number and the types of surgeries performed. Figure 3.2 shows two image pairs containing modifications based on the number of surgeries. Humans can recognize such variations in facial appearance with very low, or moderate level of difficulty. However, plastic surgeries can negatively impact the performance of automatic face recognition systems [115] because of the following reasons:

- Most face recognition algorithms take the holistic appearance of the face into account for feature extraction. A wide number of plastic surgeries can alter the overall appearance of the face, thereby reducing the similarity between genuine image pairs.

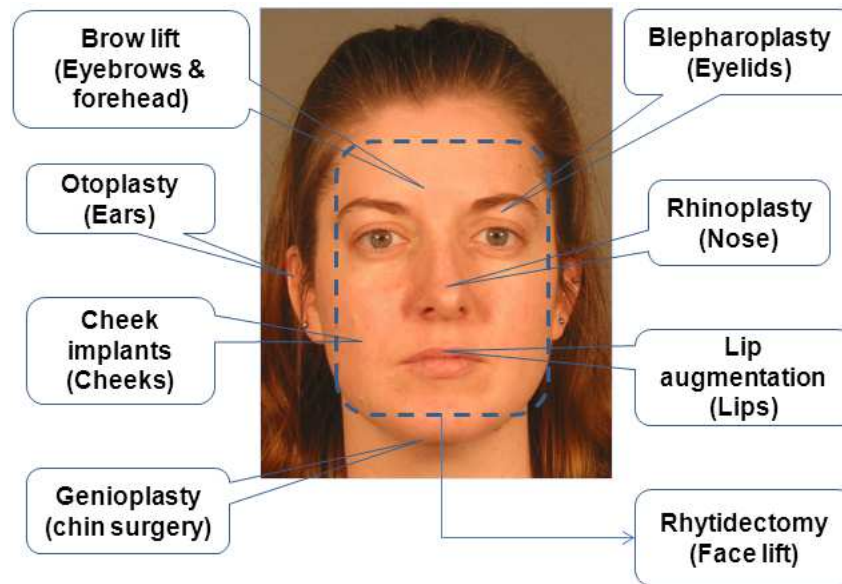


Figure 3.1: Some of the major facial plastic surgeries. Image taken from the FRGC database [13].

- Depending on the type and number of surgeries performed, a multitude of variations are possible in the appearance of the face. Such variations are difficult to be modeled by existing face recognition algorithms.

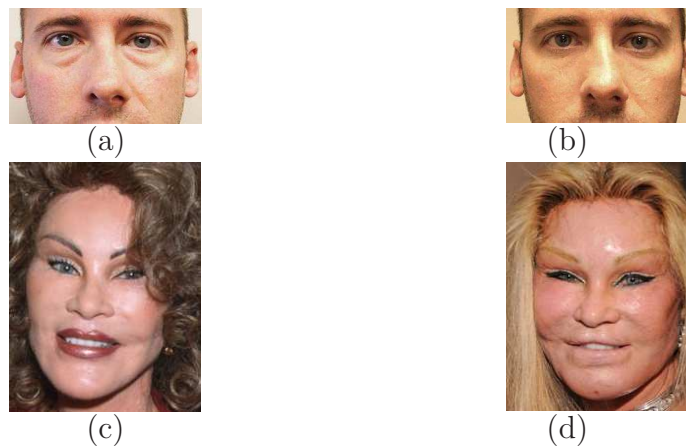


Figure 3.2: Images showing the degree to which the appearance of a human face can be modified by plastic surgeries. Top row: (a) before and (b) after a minor plastic surgery (blepharoplasty). Bottom row: (c) before, and (d) after multiple plastic surgeries.

In some cases, facial plastic surgery can unintentionally serve as a method to circumvent automatic face recognition systems. This can be a considerable security risk at locations where automatic face recognition systems are actively employed (e.g., airports).

Only recently, have researchers from the biometric community begun to investigate the effect of plastic surgery on face recognition algorithms [114,116,117]. Prior to that, research on this topic was stymied by the lack of databases containing pre- and post-surgery face images. Singh et al. [114] assembled the first database that contains face images related to various types of plastic surgeries. The low recognition accuracies that have been reported on this database seem to suggest that the task of face recognition on plastic surgery images is a challenging problem.

3.3 Existing Approaches

Singh et al. [114] reported recognition accuracies on the plastic surgery database using six different face recognition algorithms: Principal Component Analysis (PCA), Fisher Discriminant Analysis (FDA), Local Feature Analysis (LFA), Circular Local Binary Patterns (CLBP), Speeded Up Robust Features (SURF), and Neural network Architecture based 2-D Log Polar Gabor Transform (GNN). These algorithms were selected because they provide a combination of appearance-based, feature-based, descriptor-based, and texture-based feature extraction and matching approaches. Despite combining local and global recognition approaches, the matching performance obtained was rather low (see Table 3.1). Marsico et al. [118] used correlation-based face recognition on pose and illumination normalized images. Bhatt et al. [116] used an evolutionary granular approach with CLBP and SURF features to process tessellated face images. Aggarwal et al. [117] used a combination of face recognition by parts and sparse representation approach. The matching schemes used in the literature, along with their rank-one recognition accuracies are listed in Table 3.1.

3.4 Motivation

A careful study of the existing research in this area reveals the following interesting observations:

1. A majority of the algorithms that have been used are learning based which require a carefully selected set of training images. Despite this, it can be observed that the rank-one identification accuracy did not exceed 79%.
2. No commercial face recognition systems have been used for evaluating recognition

Table 3.1: List of algorithms used for performing face recognition on plastic surgery images and the corresponding rank-one accuracies.

Authors	Algorithm used	Rank-one Accuracy
Singh et al. [114]	PCA	29.1%
	FDA	32.5%
	LFA	38.6%
	CLBP	47.8%
	SURF	50.9%
	GNN	54.2%
Marsico et al. [118]	Correlation based approach	70.6%
Bhatt et al. [116]	Evolutionary granular approach	78.6%
Aggarwal et al. [117]	Combination of recognition-by-parts & sparse representation approaches	77.9%

performance.

3. No biometric *fusion* schemes have been explored in an attempt to improve recognition accuracy.

Considering the rapid advancements in the area of face recognition, there is a need to improve recognition accuracy on facial images exhibiting plastic surgeries. To this end, the present chapter provides the following contributions:

1. The recognition performance of two commercial face recognition systems on plastic surgery images is evaluated. It is demonstrated that these systems can provide performance on par with the learning based methods.
2. An information fusion approach that combines independently processed ocular information with the face biometric is presented. The proposed approach is observed to provide the current highest reported recognition performance on plastic surgery images.

The usage of ocular information for this problem provides the following benefits:

1. An empirical analysis suggests that the number of plastic surgeries that affect the appearance of the ocular region, compared to those that alter the holistic appearance of the face, is very small. Table 3.2 shows a list of surgeries categorized based on the primary facial region impacted by the surgery. It is apparent from this table that only a few of the surgeries directly impact the ocular region. Thus, in post-surgery

images, the ocular region is likely to be more stable than the global facial appearance.

Sample images demonstrating this observation are provided in Figure 3.3.

Table 3.2: **List of major facial plastic surgeries separated by the corresponding regions whose appearance can be potentially affected.**

Primary region of impact	Type of surgery
Entire face (10)	Rhinoplasty, Genioplasty, Cheek implant, Otoplasty, Liposhaving, Skin resurfacing, Rhytidectomy, Lip augmentation, Craniofacial surgery, Dermabrasion
Only the ocular region (3)	Blepharoplasty, Brow lift, Non-surgical local procedures (e.g., BOTOX)



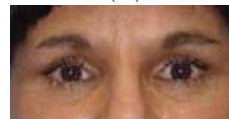
(a)



(b)



(c)



(d)

Figure 3.3: Facial images of a subject (a) before, and (b) after undergoing rhytidectomy. (c) and (d): Corresponding ocular images of the same subject. Note that the variation in the appearance of the face, from a visual perspective, is much larger than that of the ocular region.

- Existing research suggests that the fusion of ocular information with the face biometric can lead to improved recognition performance [24].

3.5 Ocular Recognition

The ocular region refers to a small region around the eye, containing the eye, the eyebrows, and the surrounding skin. Recent research has shown that the ocular information can be used as a soft biometric [24, 119]. It has been experimentally demonstrated that the ocular information can be used in lieu, or to improve the matching accuracy, of the iris [4] and face [24] under non-ideal conditions. While there are no specific guidelines for the dimensions of the ocular region, Park et al. [24] suggest that including the eyebrows can

result in higher matching accuracy. Most existing approaches use monocular information from either the left or right side of an individual's face. In this study, information corresponding to both the eyes (bi-ocular [92]) is considered. The reasons for using bi-ocular information are:

1. Park et al. [24] showed that the fusion of the left and right ocular region improves matching accuracy.
2. The spatial resolution of the face images used in this work is very low (explained in Section 3.7). Thus, utilizing the bi-ocular region ensures an effective use of information.

Some examples of the bi-ocular images used in this work are shown in Figure 3.4.



Figure 3.4: Sample bi-ocular images used in this work. Note that the images have been resized for the purpose of clarity.

3.6 Proposed Approach

Based on the initial hypothesis, the proposed approach combines the information from the face and ocular regions at score level to improve the recognition performance. Two commercial face recognition software, Verilook 3.2 [120] and PittPatt [121], were used in this work. The use of these software helps in establishing baseline performances due to commercial face recognition systems on plastic surgery images. This also helps in avoiding computationally expensive training based methods.

To perform automatic cropping of ocular regions from face images, a face detector based on the Viola-Jones Adaboost algorithm [122] was used. This step also serves as a basic quality check, where challenging images that could cause Failure To Enroll (FTE) error are discarded (e.g., images containing very small inter-ocular distances, partial faces, etc.). Ocular regions extracted from low-resolution face images could be very noisy and

impact the recognition performance. To perform feature extraction from ocular regions, two techniques, viz., Scale Invariant Feature Transform (SIFT) [123] and Local Binary Patterns (LBP) [124] were used. The combination of SIFT and LBP techniques allows for image feature extraction at both local and global levels, respectively. Furthermore, SIFT and LBP have been the most significantly used techniques* in the ocular recognition literature [4, 24]. The use of these techniques helps in maintaining uniformity for performance comparisons.

3.6.1 Scale Invariant Feature Transform

The Scale Invariant Feature Transform (SIFT) technique works by detecting and encoding information around local keypoints that are invariant to scale and orientation changes of an image. Given an image $I(x, y)$, the corresponding scale space image $L(x, y, \sigma)$, at a scale σ , is obtained as $L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$, where $G(x, y, \sigma)$ is a Gaussian filter and the symbol $*$ represents a convolution operation. A set of Difference of Gaussian (DoG) images, between scales separated by a multiplicative factor k , are obtained by the equation $DoG = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y)$. From this set of images, extrema points are detected by choosing the local maxima or minima among eight neighbors of a pixel in the current image, and nine neighbors each in the scales above and below the current DoG image. These extrema points correspond to image discontinuities and are further processed to exclude unstable extrema points. A 36 bin orientation histogram covering the $[0, 360]$ interval around each keypoint is then generated using the gradient magnitude $m(x, y)$ and orientation $\theta(x, y)$ information, where $m(x, y) = [((L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2)]^{\frac{1}{2}}$, and $\theta(x, y) = \tan^{-1} \left(\frac{(L(x, y+1) - L(x, y-1))}{(L(x+1, y) - L(x-1, y))} \right)$. The orientation of the keypoint is computed as the highest peak in the orientation histogram associated with it. The feature vector is obtained by sampling the gradient magnitude and orientations within a descriptor window of size 16×16 around a keypoint. The final keypoint descriptor of dimension $4 \times 4 \times 8$ is generated by computing an 8 bin orientation histogram over 4×4 sample regions within the descriptor window. In this work, a publicly available MATLAB implementation [125] of SIFT was used.

*Gradient Orientation Histogram (GO), another global level feature extraction technique, has also been widely used in ocular recognition literature. However, it was excluded in this study because LBP outperformed GO.

3.6.2 Local Binary Patterns

Given an image I , sample points are first determined by uniformly sampling the image at a fixed frequency. A block of size 8×8 pixels around every sampling point is considered as a region of interest (ROI). For each pixel p within the ROI, a neighborhood of size 3×3 pixels is considered for LBP value generation, as shown in Figure 3.5.

p_1	p_2	p_3
p_0	p	p_4
p_7	p_6	p_5

Figure 3.5: Neighborhood for computing the LBP of pixel p .

The mathematical equation for computing the LBP value at a pixel p is given by:

$$LBP(p) = \sum_{k=0}^{k=7} 2^k f(I(p) - I(p_k)), \quad (3.1)$$

where $I(p_k)$ represents the intensity value of pixel p_k , and

$$f(x) = \begin{cases} 1 & \text{if } x \geq 0, \\ 0 & \text{if } x < 0. \end{cases} \quad (3.2)$$

The LBP values of all the pixels within a given ROI are then quantized into an 8 bin histogram. Histograms corresponding to all sampling points are then concatenated to form a final feature vector. Euclidean distance was used to measure the similarity between two feature vectors. In this work, to perform LBP feature extraction and matching, every *RGB* ocular image was first decomposed into its individual *R*, *G*, and *B* channels. Each channel was sampled at a frequency of 16 pixels, yielding a total of 465 sampling points. The final LBP feature vectors for each channel were of size 1×3720 (concatenating 8 bin histograms for 465 sampling points).

3.6.3 Score-level Fusion

For a given image, let S_{VL} and S_{PP} denote the face match scores obtained using Verilook and PittPatt, respectively. S_{SIFT} , represents the SIFT ocular score and S_{LBP-R} , S_{LBP-G} , and S_{LBP-B} represent the LBP ocular scores for each of the *R*, *G*, and *B* channels of an ocular image, respectively. A final LBP ocular score, S_{LBP} , was computed by considering the average of S_{LBP-R} , S_{LBP-G} , and S_{LBP-B} . The averaging operation was chosen because

it provided relatively better performance, when compared to the other operators (e.g., min, max, etc.). Score-level fusion was then performed to combine the face and ocular information. A schematic representation of the proposed score-level fusion approach is shown in Figure 3.6.

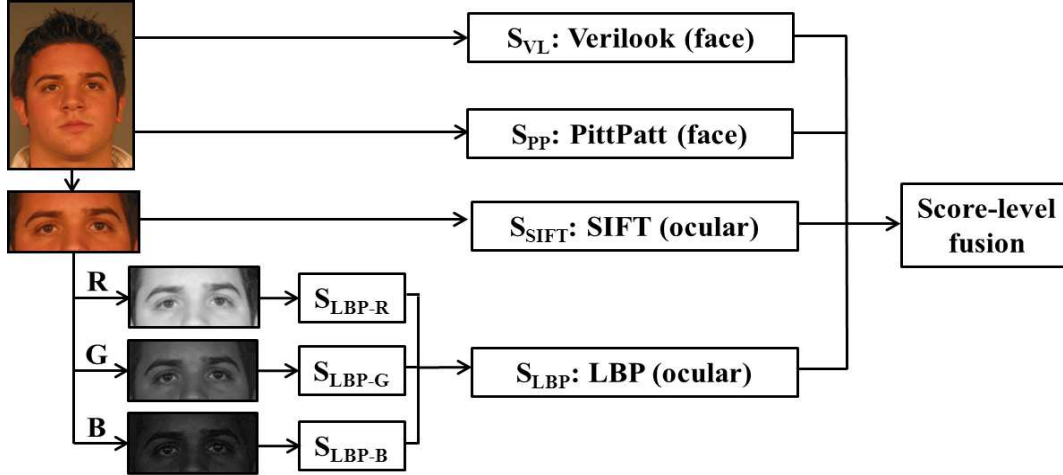


Figure 3.6: A schematic representation of the proposed approach.

3.7 Database

Images from the plastic surgery database described in [114] are used in this work. Currently, this is the only publicly available database that contains images of subjects captured before and after various types of plastic surgeries. Biometric databases are typically assembled through a concerted data collection process by acquiring the required data from the subjects directly. On the contrary, this database was generated by downloading facial images from two different plastic surgery information websites[†]. This introduces significant challenges in working with this database, such as: (a) low resolution, (b) variations in scale and expression, and (c) duplicate entries. Figure 3.7 shows sample images illustrating these challenges.

Three different datasets are considered in this work. The details of each dataset are listed as follows:

[†]www.locateadoc.com and www.surgery.org

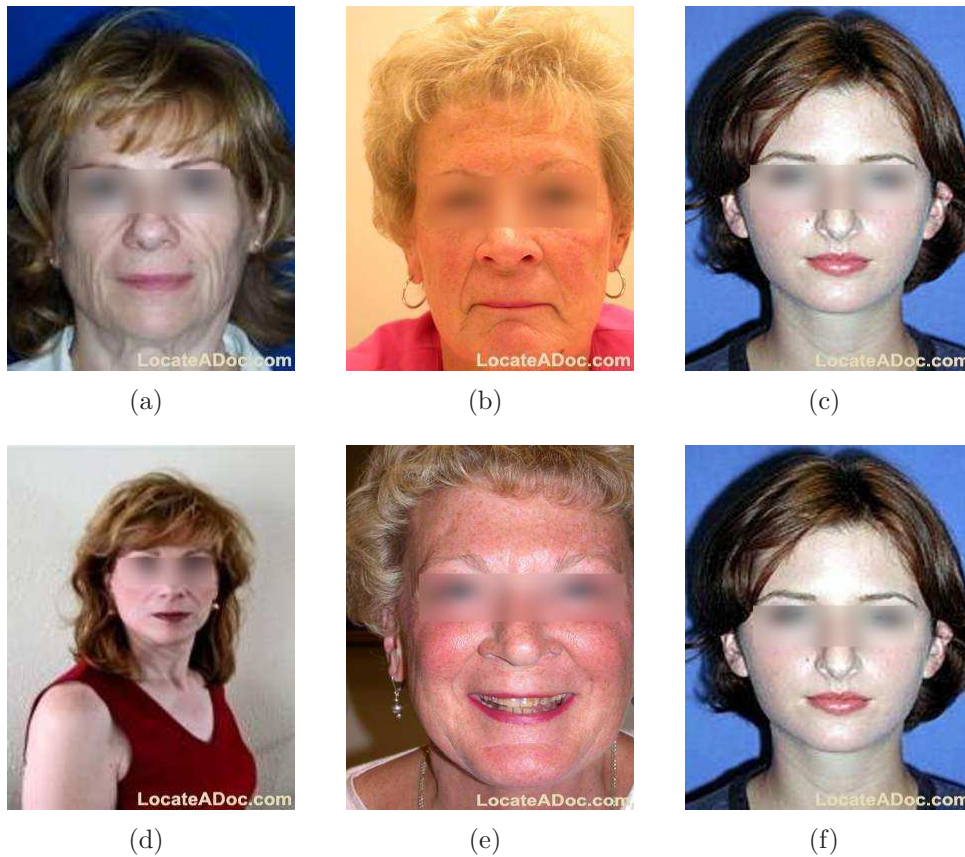


Figure 3.7: Images exhibiting some of the challenges in the facial plastic surgery database. (a) and (d): images with varying resolution, scale and inter-ocular distances corresponding to the same subject. (b) and (e): variations in expressions of a subject. (c) and (f): duplicate entries. The image in (c) is listed as ID #26300 and its duplicate image in (f) is re-listed as ID #28519. Note the difference in identification labels, although they belong to the same subject who has undergone multiple surgeries. This incorrect labeling can negatively impact the perceived matching accuracy.

3.7.1 Face Dataset A

All the images contained in the plastic surgery database were used in this dataset. This dataset contains frontal face images of 900 subjects. For each subject, there is 1 pre-surgery facial image and 1 post-surgery facial image. The resolution of the images range from 163×131 to 288×496 pixels, and the inter-ocular distance varies from 20 to 100 pixels. These images are divided into a gallery (containing 900 pre-surgery images), and a probe set (containing the corresponding 900 post-surgery images). This dataset helps in performing a direct comparison of recognition performances obtained by commercial recognition systems, with those reported in the existing literature.

3.7.2 Face Dataset B

This dataset was obtained by discarding images from *face dataset A* corresponding to: (a) failures in face detection using the Adaboost algorithm, and (b) very low image resolution that can yield noisy ocular regions (as described in Section 3.6). As a result, a total of 478 images corresponding to 239 subjects were selectively discarded from *face dataset A*. The remaining 1322 images are divided into a gallery (containing 661 pre-surgery images), and a probe set (containing the corresponding 661 post-surgery images). A set of 568 face images corresponding to 568 unique subjects from the FRGC database [13] were added to the gallery. These images have a resolution of 1704×2272 pixels, with an average inter-ocular distance of 260 pixels. These additional images help in (a) compensating for the effect of discarded images, (b) observing the robustness of the proposed feature extraction and matching techniques by increasing the number of impostor scores, and (c) providing a heterogenous combination of surgically modified and unmodified face images.

3.7.3 Ocular Dataset

This dataset was generated by automatically cropping the bi-ocular regions from images in *face image dataset B*. The average resolutions of the cropped bi-ocular regions range from 115×54 to 842×392 pixels. All the ocular images in both the gallery and probe sets were resized to a fixed resolution of 500×250 pixels. This helps in ensuring a fixed-size feature vector when global feature extraction schemes are used.

The total number of images used in the face and ocular datasets, along with their spatial resolutions are summarized in Table 3.3.

Table 3.3: Number of images used in each dataset, along with their spatial resolutions.

		Number of images	Resolution (in pixels)
Face dataset A	Gallery	900	163×131 to 334×466
	Probe	900	147×226 to 288×496
Face dataset B	Gallery	1229 (661 + 568)	288×250 to 1704×2272
	Probe	661	288×250 to 288×485
Ocular dataset	Gallery	1229	500×250
	Probe	661	500×250

3.8 Experiments and Results

To determine the face recognition performance, every image in the probe set of the face image dataset was matched against the gallery. The same protocol was used for the ocular image dataset to generate the ocular match scores. When performing score-level fusion, the score matrices corresponding to *face dataset B* and *ocular dataset* were normalized in the $[0, 1]$ range using min-max normalization.

3.8.1 Face recognition performance

Cumulative Match Characteristic (CMC) curves were used to summarize the obtained identification performances. Figure 3.8 shows the CMC curves obtained using the commercial face recognition systems on the considered face datasets. The rank-one recognition accuracies obtained using Verilook and PittPatt on *face dataset A* were observed [‡] to be (a) 70.3% (b) 65.8%, respectively. Similarly, the corresponding rank-one recognition accuracies obtained on *face dataset B* were observed to be (a) 73.9% and (b) 81.4%, respectively. From the figure, it can be observed that PittPatt provides better recognition performance than VeriLook when low resolution images are discarded.

3.8.2 Ocular recognition performance

The rank-one accuracies obtained using LBP and SIFT on the *ocular database* were observed to be 45.6% and 48.1%, respectively. The CMC curves for both the techniques are shown in Figure 3.9. From the figure, it can be observed that SIFT provides better ocular recognition performance compared to LBP. This is because SIFT depends on local key-point information that is scale and rotation invariant. On the other hand, the LBP match score is dependent on the similarity of global level information that is affected by misalignment of gallery and probe images.

3.8.3 Score-level fusion performance

Weighted score-level fusion is used to combine the normalized scores from the following scenarios: (a) face (VeriLook and PittPatt scores obtained using *face dataset B*), (b) ocular

[‡]Only these recognition performances should be considered when making a direct comparison with results from existing literature.

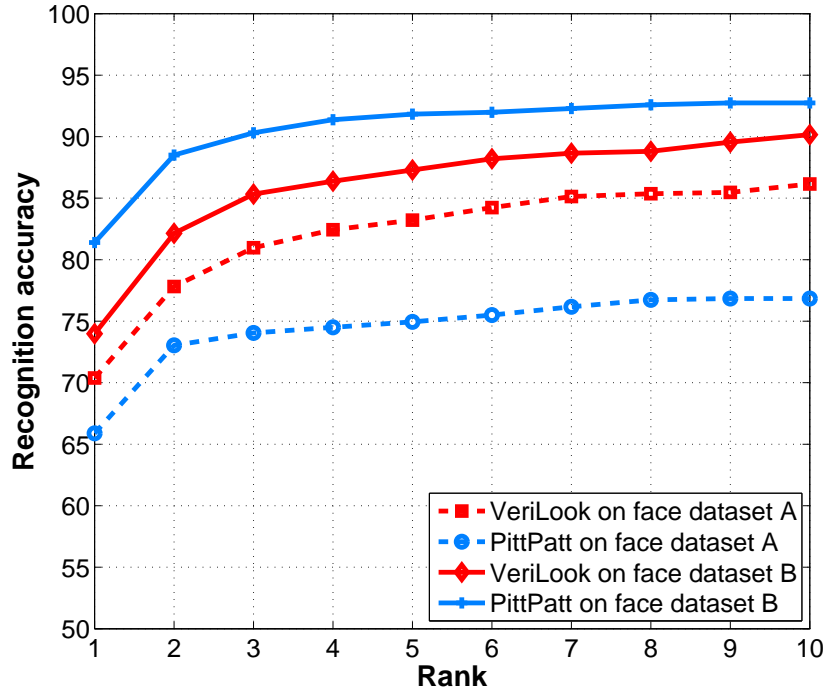


Figure 3.8: CMC curves showing the recognition performances of VeriLook and PittPatt on *face dataset A* and *face dataset B*.

(LBP and SIFT scores obtained using *ocular dataset*), and (c) face and ocular (VeriLook and PittPatt scores obtained using *face dataset B*, LBP and SIFT scores obtained using *ocular dataset*). These normalized scores were combined using the simple sum rule with different weights, with an objective of maximizing the rank-one accuracy. The rank-one recognition accuracies obtained for the above mentioned scenarios are: (a) 85.3% , (b) 63.9% , and (c) 87.4%. Figure 3.10 shows the corresponding CMC curves, along with the weights used for fusion in each case. From the results obtained, it can be observed that score-level fusion clearly improves the recognition performances when combining both inter-modality scores and intra-modality scores. The rank-one recognition performance obtained by the proposed approach (87.4%) reflects the highest recognition accuracy observed in the literature for this database. The rank-two recognition accuracy for the fusion scheme is observed to be 94.4%. This significant increase ($\sim 7\%$) in performance was due to the presence of duplicate entries, as described in Section 3.7. In such cases, a probe image would first match with the duplicate sample of the same subject (with different identification tag), and then with the the corresponding sample with the same identification tag. Such an

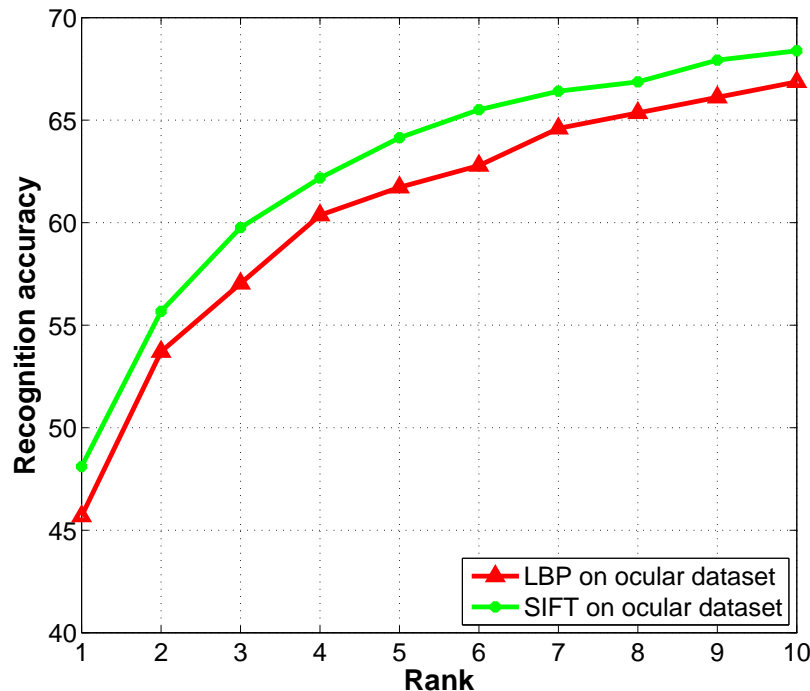


Figure 3.9: CMC curves showing the recognition performances of LBP and SIFT on *ocular dataset*.

effect causes a reduction in performance at rank-one. Some of the duplicate images that match at rank-two but not at rank-one are shown in Table 3.4. If such duplicate images are accounted for (either removed, or given the same identification tags), a higher rank-one recognition performance can be expected. The benefit of the proposed technique can be observed in Table 3.5, showing example face and ocular images that were not correctly matched at rank-1 by the face recognition systems, but were correctly matched at rank-1 after performing fusion.

3.8.4 Effect of individual surgeries

The effect of individual surgeries on the recognition performances was studied. Depending on the type of surgery performed, the images were categorized into two main groups: global and local [114]. Images corresponding to global surgeries show variations in the overall appearance of the face (e.g., rhytidectomy). Local surgeries, however, typically modify the appearance of a single facial feature, and may minimally impact the overall appearance of the face (e.g., otoplasty, rhinoplasty, etc.).

Table 3.4: Duplicate image pairs that reduce the recognition performance at rank-one. Notice the difference in the identification tags, that causes the genuine pairs to be reckoned as impostors.







Input probe image	Corresponding gallery image that the probe has to match with	Instead matches with
 <i>ID # 03918 (after)</i>	 <i>ID # 03918 (before)</i>	 <i>ID # 13176 (before)</i>
 <i>ID # 22517 (after)</i>	 <i>ID # 22517 (before)</i>	 <i>ID # 10228 (before)</i>

Table 3.5: Example face and ocular image pairs (pre- and post-surgery) that were not correctly matched at rank-1 by the face recognition systems, but were correctly matched at rank-1 after performing fusion.



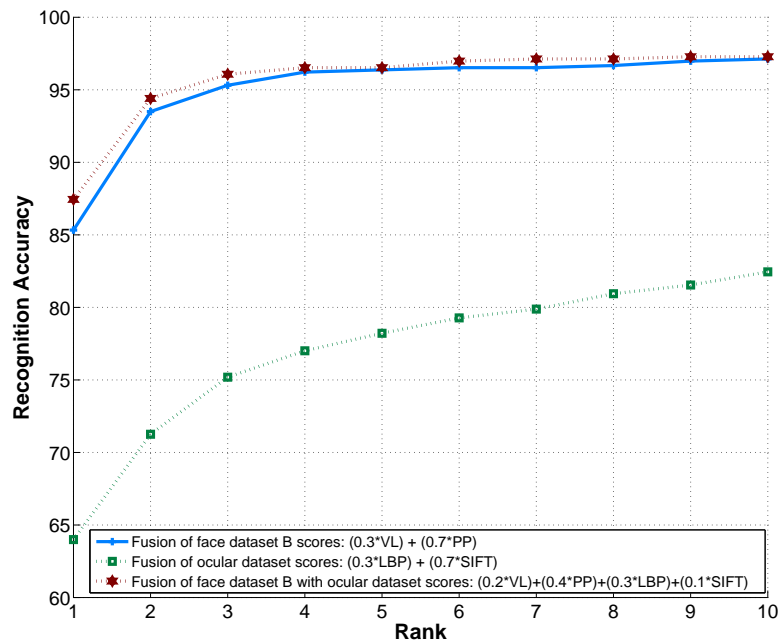


Figure 3.10: CMC curve showing the recognition accuracies obtained using score-level fusion of face scores, ocular scores, and a combination of the two.

For this experiment, images corresponding to only major surgeries are considered. Images related to surgeries that do not provide clear information about which facial region they affect were excluded. For example, botox injections can be used to modify both local (say, around the lips), as well as the global appearance. Since the database does not provide meta-data that clearly explains these details, such images were excluded from this experiment. The rank-one recognition accuracies corresponding to individual surgeries obtained using face, ocular, and fusion schemes are provided in Table 3.6.

Table 3.6: **Rank-one recognition accuracies corresponding to individual surgeries obtained using the face, ocular, and fusion schemes on images from face dataset B and the ocular dataset.**

Type of surgery	Face		Ocular		Proposed
	(VL)	(PP)	(SIFT)	(LBP)	
Browlift	88.2%	100%	64.7%	58.8%	97.0%
Otoplasty	85.4%	90.9%	69.0%	65.4%	94.5%
Blepharoplasty	74.2%	92.8%	64.2%	45.7%	94.2%
Rhinoplasty	79.1%	85.9%	54.3%	54.3%	85.9%
Rhytidectomy	78.8%	90.0%	48.4%	46.7%	92.2%

From the table, it can be observed that PittPatt and SIFT provide comparatively better

face and ocular recognition performances, respectively. Once again, the proposed approach improves the recognition performance compared to individual techniques. Singh et al. [114] performed a similar study and concluded that face recognition algorithms cannot handle global facial plastic surgeries. Similar observation can be made from the results in this work. The recognition performance is more negatively impacted by global surgeries (rhinoplasty and rhytidectomy) than local surgeries (browlift, otoplasty, and blepharoplasty).

3.9 Summary

This chapter describes a fusion approach that combines the face and ocular information to improve biometric identification using images corresponding to facial plastic surgeries. The proposed approach yields a rank-one recognition accuracy of 87.4%, which quickly increases to 94.4% at rank-two. The performance obtained using the proposed approach reflects the current best rank-one accuracy reported on the considered plastic surgery database. Compared to existing approaches, the proposed scheme presents a method to improve recognition performance without using training-based methods. Based on the results, it is opined that the problem of face recognition using the publicly available plastic surgery database could be further improved if the non-ideal factors (e.g., duplicate entries, low image resolutions, etc.) of the database are accounted for.

Chapter 4

Ocular Matching in Face and Iris Images

4.1 Introduction

One of the significant challenges in biometric recognition is the process of matching or fusing information obtained from multiple sources. Such sources of information often vary with respect to the biometric modalities, sensors, samples, or matching algorithms used [126]. Current research in biometrics is geared towards mitigating variance in information prior to its usage. In the realm of iris biometrics, Chapter 2 of this thesis proposes a technique to improve recognition performance by effectively fusing information obtained from *multiple samples* of a subject. In this chapter, the problem of matching ocular information from images corresponding to *multiple modalities*, viz., face and iris, is considered. Face and iris images are typically acquired in the visible and near-infrared spectrum of light, respectively. This presents a challenging problem for biometric matching using *cross-modality*, *cross-wavelength*, *cross-sensor*, and *cross-resolution* images. Sample *RGB* face and the corresponding *NIR* iris images are provided in Figure 4.1. A close up of the ocular regions within the considered images is provided in Figure 4.2.

Owing to their common presence in both face and iris images, the current work proposes the usage of iris and ocular region information to perform matching. This chapter aims to address the following questions:

1. Can the iris information extracted from *RGB* face and *NIR* iris image be effectively matched?
2. What is the role of ocular region information in improving the recognition perfor-

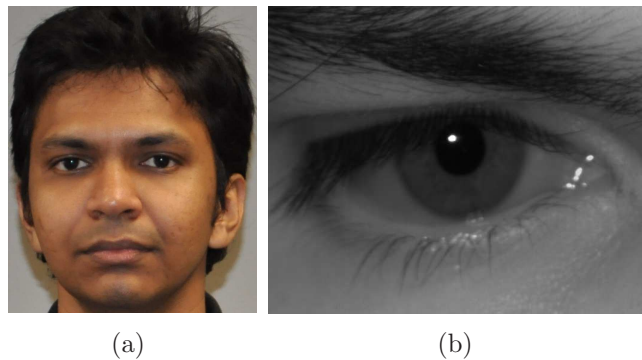


Figure 4.1: Sample images corresponding to (a) face, and (b) iris modalities, respectively.

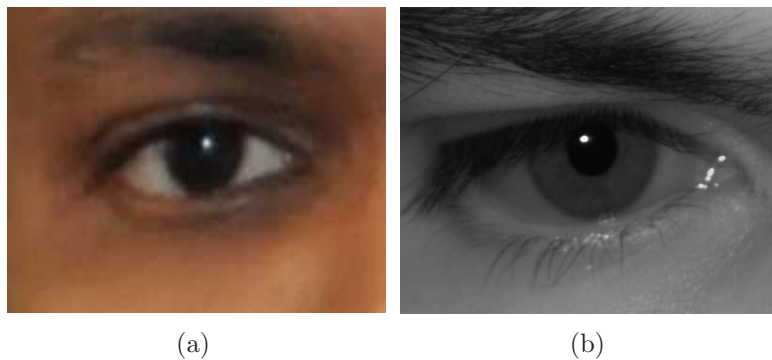


Figure 4.2: Closeup of the left-side ocular regions from (a) face, and (b) iris images shown in Figure 4.1. The goal is to perform ocular matching using the regions observed in (a) and (b). Notice the variations in scale, resolution, image acquisition wavelength, viewing angle, and the level of detail. The average number of pixels across the irides in (a) and (b) are 35 and 110 pixels, respectively.

mance?

3. What type of challenges are encountered when matching the ocular information in face and iris images? How can they be mitigated?

4.1.1 Motivation

Ocular matching in face and iris images can be required when the acquisition or the recognition stages are limited by: (a) modality of images available, and (b) type of sensor used. Such requirements are typically encountered in the following scenarios:

- *Matching legacy databases*: Given the growing interest in biometric recognition, it is increasingly possible to encounter situations when multiple databases corresponding to the identities of same individuals are merged. In such situations, the biometric modalities available in the independent databases may not always be the same. An

example of such a situation is illustrated in Figure 4.3. The process of reliably associating the identities between constituent databases can be complicated if: (a) the meta-data corresponding to the images in the individual databases are not comparable, and (b) the organizations maintaining the databases do not allow complete meta-data sharing. In such cases, cross-modality ocular matching in face and iris images can be very useful.

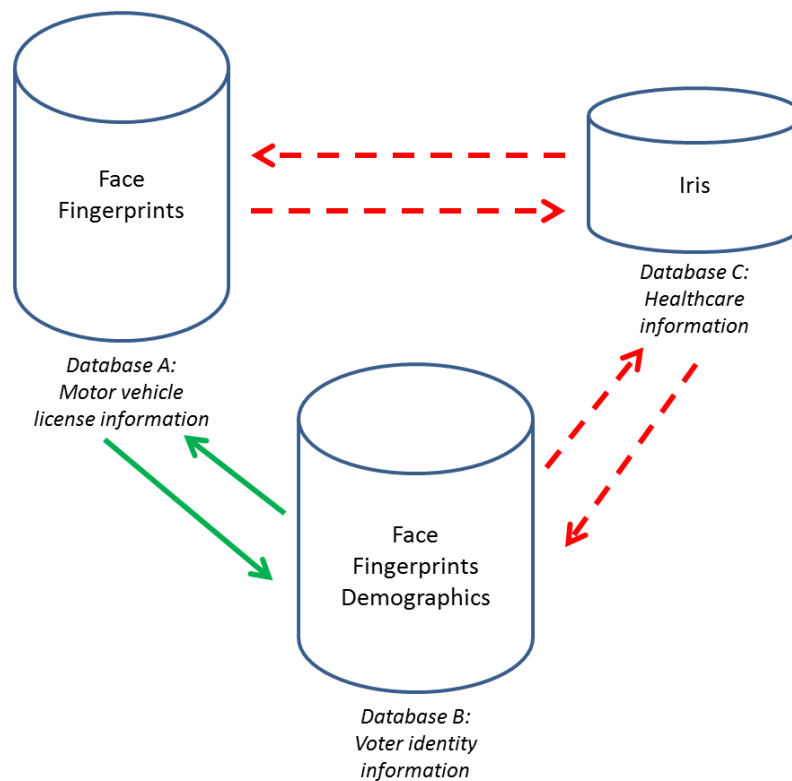


Figure 4.3: An illustration depicting the different modalities in databases. Note that relating the identities stored in *Database C* with those stored in other databases can be a challenging process.

- *Surveillance and law enforcement:* In various law enforcement scenarios, it is possible that a surveillance image has to be reliably matched with the entries in a watch-list. When a false non-match occurs due to a single modality, cross-modality matching could be beneficial. An example of such a situation is shown in Figure 4.4. When a face image acquired using a surveillance camera retrieves no match within the face database, ocular regions could be matched with an iris database.

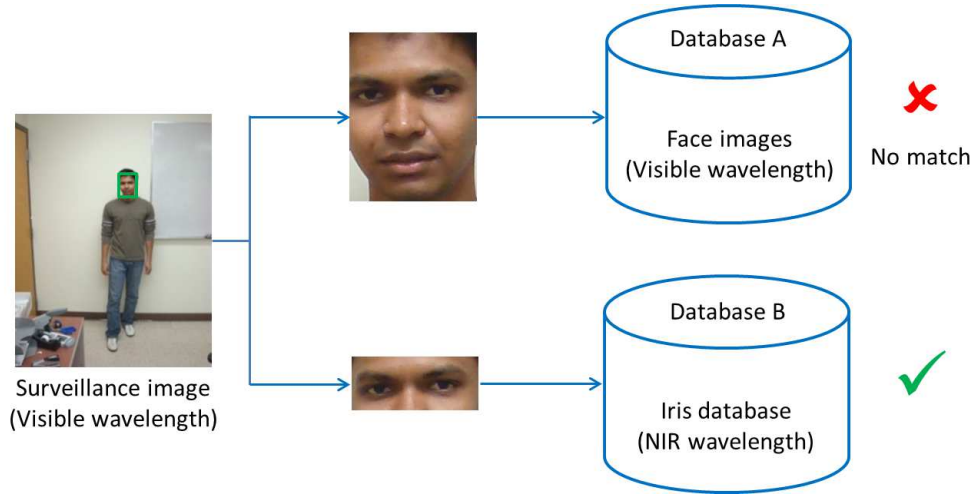


Figure 4.4: A sample scenario depicting the need for ocular matching in face-iris image pairs.

4.2 Problem Significance

Currently, the problem of matching ocular regions in face images to those in iris images (and vice-versa) has not been addressed in the literature. Some researchers have already studied iris and ocular region matching under variations in image resolution [127], [91], or variations in imaging wavelength [56]. However, the current work is significantly different from such efforts because of the *collective* consideration of the following factors:

1. *Cross-modality*: Images corresponding to both face and iris traits are used in this work.
2. *Cross-wavelength and cross-sensor**: Face images are typically acquired using sensors that operate in visible wavelength. On the other hand, iris images are acquired in the near-infrared spectrum.
3. *Cross-resolution*: Owing to the variation in sensors used, ocular regions in face images tend to be of lower resolution than those in iris images.

Table 4.1 lists the publications studying the impact of one or more of the above mentioned factors on ocular recognition performance. The significance of the current work can be

*Note that cross-sensor does not always mean cross-wavelength. For example, cross-sensor iris recognition refers to the task of matching iris images acquired using different sensors, all operating in the near-infrared spectrum.

clearly seen from the table[†].

Table 4.1: **Publications studying the impact of varying imaging factors on iris and ocular recognition (listed chronologically).**

(Year) Authors	Iris or Ocular	Images matching			
		Cross resolution or stand-off distance	Cross-sensor	Cross-wavelength	Cross-modality
(2009) Jillela and Ross [127]	Iris	Yes	No	No	No
(2009) Ross et al. [56]	Iris	No	No	Yes	No
(2010) Bharadwaj et al. [91]	Ocular	Yes	No	No	No
(2011) Connaughton et al. [128]	Iris	No	Yes	No	No
(2012) Xiao et al. [129]	Both	Yes	Yes	No	No
(2013) Tan and Kumar [130]	Both	Yes	No	No	No
Current work	Both	Yes	Yes	Yes	Yes

4.3 Database

Face and iris images from the Biometric Collection of People (BioCoP) database [19] were used in this work[‡]. Both the face and iris images were acquired in two different sessions (viz., *Set 1* and *Set 2*). *Set 1* and *Set 2* contain images corresponding to 704 and 654 subjects, respectively. All the 654 subjects of *Set 2* overlap with those in *Set 1*. Both sets contain 1 face, and 2 iris images (corresponding to the left and right sides) of a subject. Face images were acquired using a Olympus C-8080 wide zoom camera, operating in the visible spectrum with a resolution of 2448×3264 pixels. Iris images were acquired using an Oki IrisPass M sensor in the NIR spectrum, with a resolution of 640×480 pixels. The average radius of the iris in the face and iris images was observed to be 35 pixels and 110 pixels, respectively. The variation in iris radii and spatial resolutions occur due to different

[†]Note that the table lists only those publications which study the impact on recognition performance by varying one or more imaging factors. There is a considerable amount of work on iris segmentation in visible wavelength (e.g., [78]). However, such publications are not listed in the table as they do not focus on recognition performance. Multiple publications exist on some topics such as cross-sensor iris recognition. Only the earliest of such attempts are listed.

[‡]A copy of the BioCoP database release agreement (consent form) is provided in Appendix B

stand-off distances for the two sensors. The stand-off distances for face and iris images was maintained to be 2 meters and 30-60 centimeters, respectively. Ocular regions of size 225×169 pixels (approximately) were manually cropped from the face images. To ensure uniformity in the region of interest (ROI) between iris and ocular images, the following two criteria were imposed:

1. Center of the iris was maintained as the center of ocular region image.
2. The aspect ratio of the ocular region from the face image was maintained to be the same as that of the iris image.

A summary of the database specifications is provided in Table 4.2. Sample images of the face, cropped ocular region, and the corresponding iris are shown in Figure 4.5.

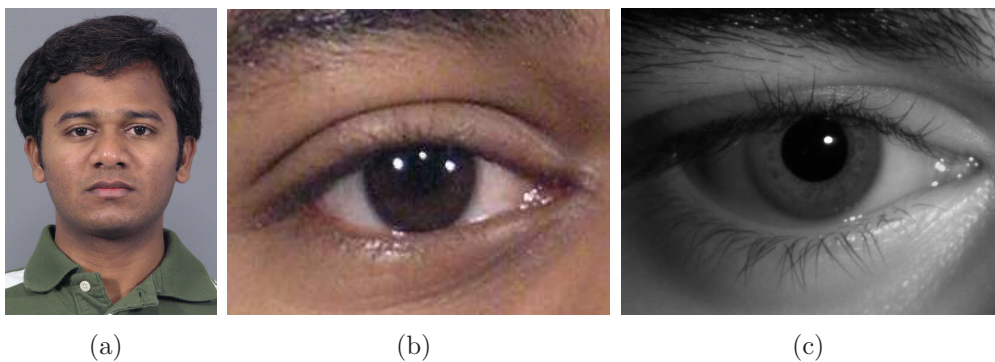


Figure 4.5: Sample images from the BioCoP database showing the (a) face, (b) cropped ocular region, and (c) iris of a subject. *Subject's approval to use the images for illustration purposes is on file.*

Table 4.2: Specifications of the considered BioCoP database [19].

Modalities	Sensor used	Imaging spectrum	Image resolution (pixels)	# of subjects		Average iris radius (pixels)	Stand-off distance
				Set 1	Set 2		
Iris	Oki IrisPass M	NIR	640×480	704	654	110	30-60 cm
Face	Olympus C-8080 wide zoom	VIS	2448×3264	704	654	35	2m
Ocular (cropped from face)	Olympus C-8080 wide zoom	VIS	2448×3264	704	654	35	2m

4.3.1 Challenges

Variations in the following factors are considered to render the database challenging:

- **Sensors and imaging wavelengths:** The appearance of iris texture in face and iris images is significantly different due to variations in sensors, and imaging wavelengths (see Figure 4.6).
- **Iris radius:** The stand-off distances and resolution of the face and iris sensors were different. This causes a difference in the radii of the irides within the corresponding pair of face and iris images. Figure 4.6 illustrates this effect.

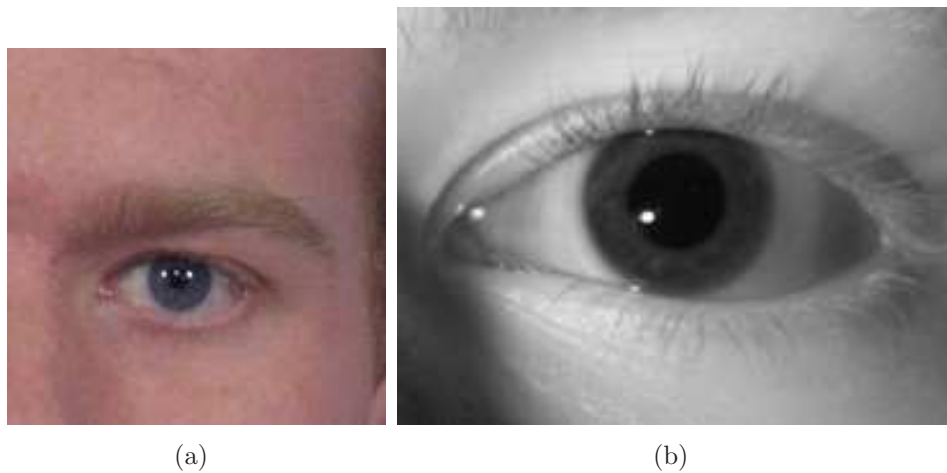


Figure 4.6: Corresponding ocular regions from (a) a face image and (b) an iris image acquired under visible and NIR spectra, respectively. Notice the variation in textural appearance of the iris within the images.

- **Viewing angle:** The face and iris sensors were placed at different heights from the ground level, causing appearance variations within the corresponding pairs of ocular regions. Figures 4.7(a) and (b) illustrate the described effect. Notice that the folds between the upper eyelid and the eyebrow seen in (a), do not have the same appearance in (b).
- **Illumination:** Images obtained by the iris sensors were observed to exhibit significant illumination variations, as shown in Figure 4.8. Such variations render the task of iris segmentation and ocular feature extraction very challenging.



Figure 4.7: Images showing variations in the viewing angle between corresponding ocular regions obtained from (a) a face image and (b) an iris image, respectively.

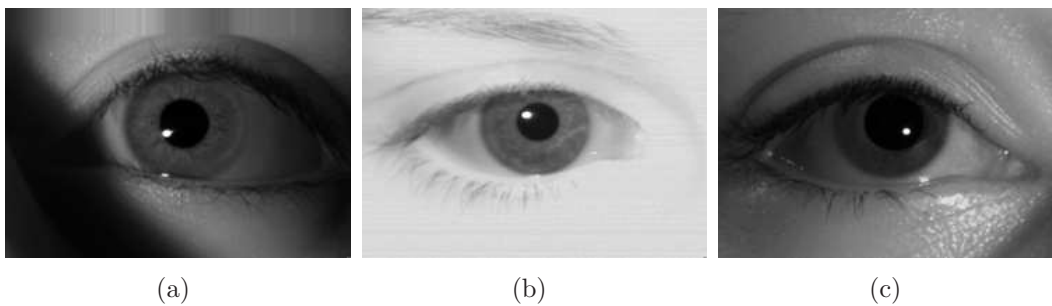


Figure 4.8: Variations in illumination observed in images acquired by the NIR iris sensor.

- **Occlusions:** A large number of images were observed to contain occlusions of the iris and ocular regions, caused by the eyelids, eyelashes and the hair. Sample images showing such occlusions are provided in Figure 4.9.

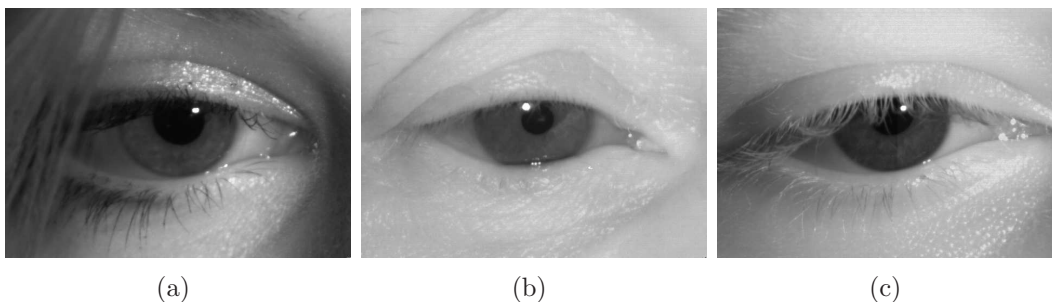


Figure 4.9: Occlusions of the iris and ocular regions caused by (a) eyelids, (b) eyelashes and (c) hair, as observed in the images acquired by the iris sensor.

- **Sensor-noise and non-uniformity in acquisition:** The Oki IrisPass M sensor depended on its in-built automatic face detection output for localizing and imaging the iris regions. Errors in such process resulted in non-uniform imaging. Furthermore, sensor noise was also observed in some images. Sample images of such cases are

provided in Figure 4.10.

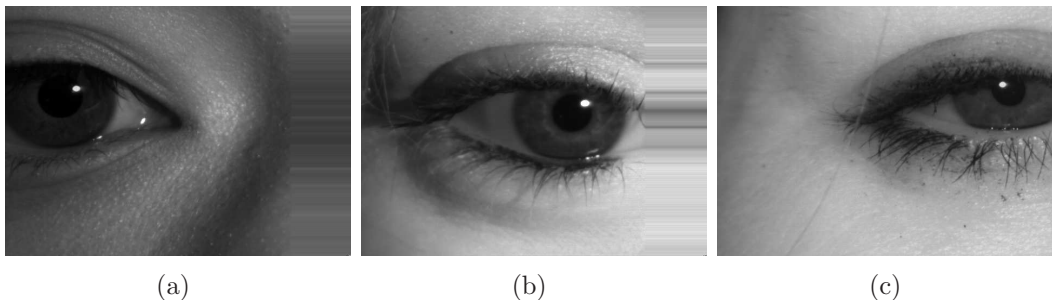


Figure 4.10: Images depicting the sensor noise and non-uniformity in imaging.

4.4 Outline of Experiments

Given the NIR iris and visible (VIS) ocular images, the following six different matching scenarios are possible:

1. Iris matching - NIR iris images with NIR iris images
2. Iris matching - VIS ocular images with VIS ocular images
3. Iris matching - VIS ocular images with NIR iris images (cross-modality)
4. Ocular region matching - NIR iris images with NIR iris images
5. Ocular region matching - VIS ocular images with VIS ocular images
6. Ocular region matching - VIS ocular images with NIR iris images (cross-modality)

The goal of this chapter is to develop techniques that can reliably perform cross-modality matching (i.e., cases 3 and 6, respectively). Section 4.5 deals with iris recognition experiments using open-source and commercial algorithms. Ocular region matching techniques are discussed in Section 4.6.

4.5 Iris Recognition

Two separate iris recognition algorithms were considered in this work:

1. *Open-source*: Libor Masek's [77] implementation of Daugman's algorithm [62], and

2. *Commercial*: VeriEye iris recognition system from Neurotechnology [105].

It was expected that VeriEye would provide better performance than Libor-Masek's implementation. This is because the latter is a rudimentary implementation of Daugman's algorithm. Utilizing both the algorithms allows for a comparison of the widely popular open-source implementation with one of the many available commercial systems. Compared to the commercial system, the open-source implementation allows more control on various factors that impact iris recognition performance (e.g., segmentation methods, feature template size, etc.).

4.5.1 Open Source Algorithm

Iris Segmentation

Libor Masek's implementation utilizes Hough transforms to perform iris segmentation [51]. Two additional iris segmentation algorithms based on Integro-Differential Operators [62] and Geodesic Active Contours [67] were also used. All the three techniques were tested on a sample set of 100 iris images and their corresponding 100 ocular images selected from the BioCoP database. The purpose of this experiment is to choose the best performing iris segmentation algorithm that can be used with Libor Masek's feature extraction and matching scheme. The segmentation accuracies [§] of all the three techniques were observed to be around 76% [¶]. The main reasons for poor segmentation performance were observed to be:

1. non-uniform illumination and occlusions in NIR iris images, and
2. low resolution and presence of dark colored irides in VIS ocular images.

The pupillary boundary in a dark colored VIS image is often difficult to distinguish, even for a human expert. Sample images showing correct and incorrect segmentation outputs obtained using the considered algorithms are shown in Figure 4.11 and 4.12, respectively.

The poor segmentation performances highlight the need for robust segmentation algorithms that can operate on both NIR and VIS images. As the focus of this work is on

[§]Segmentation accuracy = $\frac{\text{Number of correctly segmented images}}{\text{Number of input images provided}} \times 100$

[¶]Reported using the considered sample set containing 100 NIR iris, and 100 VIS ocular images

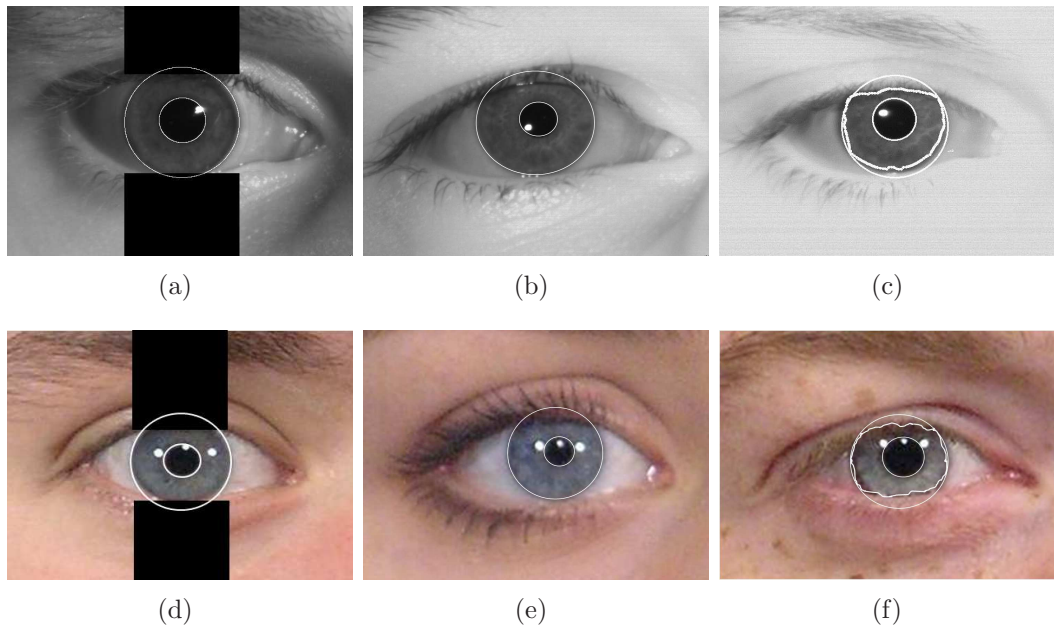


Figure 4.11: Sample NIR iris (top row) and VIS ocular region (bottom row) images showing correct iris segmentation output obtained using: Integro-Differential Operator [(a) and (d)], Hough transform [(b) and (e)], and Geodesic Active Contours [(c) and (f)] based algorithms.

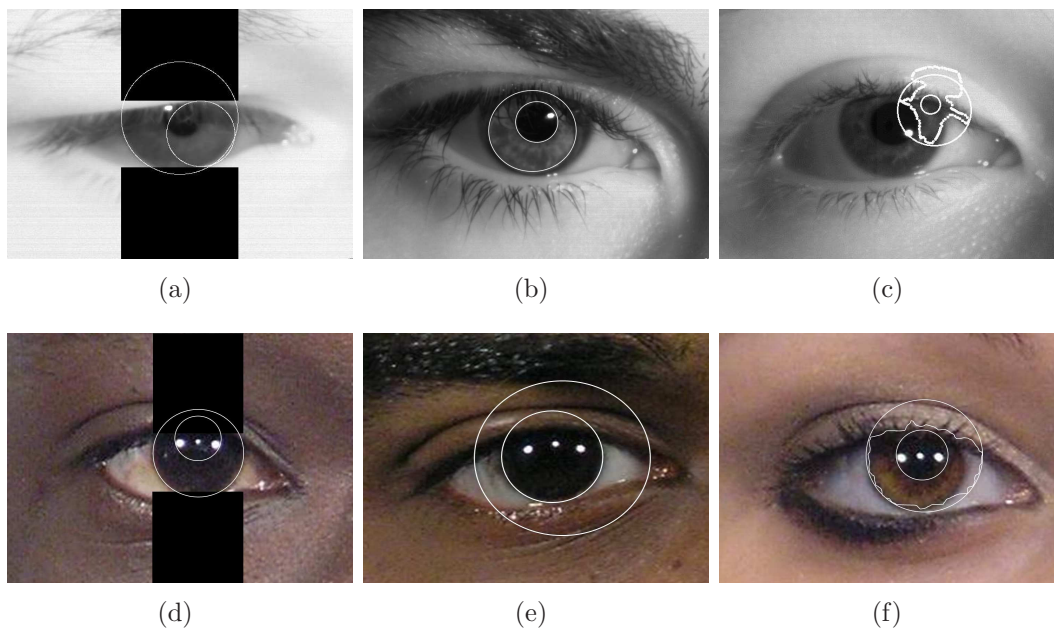


Figure 4.12: Sample NIR iris (top row) and VIS ocular region (bottom row) images showing incorrect iris segmentation output obtained using: Integro-Differential Operator [(a) and (d)], Hough transform [(b) and (e)], and Geodesic Active Contours [(c) and (f)] based algorithms, respectively..

matching, and not on segmentation, iris regions were manually segmented for further analysis. This process helps in having a reasonably reliable ground truth, while minimizing the

impact of incorrect segmentation on the recognition performance. As manual segmentation is a time consuming process^{||}, the open source algorithm was tested only on the sample set containing 100 subjects. Using 1 sample per subject does not generate genuine scores for intra-modality comparison. However, this experiment allows in observing the following aspects of iris recognition:

1. *Imaging wavelength*: Boyce et al. [55] suggest that cross-spectral iris matching performance depends on the difference of imaging wavelengths considered. In this regard, iris regions extracted from the three separate channels of the VIS ocular images, viz., R, G, and B, were individually matched against those extracted from the NIR iris images. An NIR iris image, along with the corresponding R, G, and B channel images extracted from a VIS ocular image are shown in Figure 4.13.

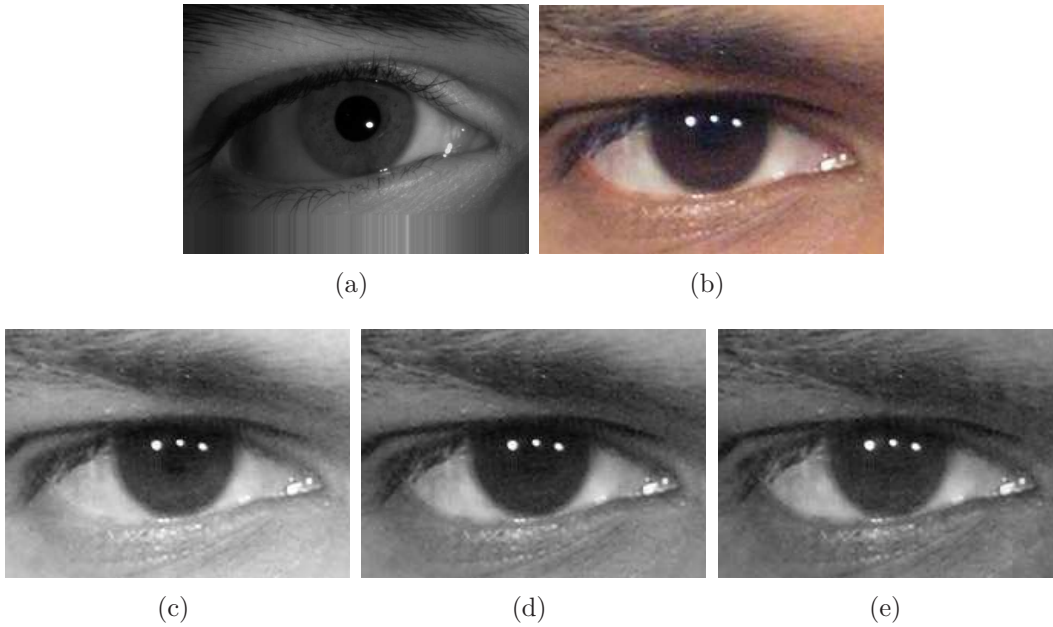


Figure 4.13: (a) Sample *NIR* iris image. (b) Corresponding *VIS* ocular region cropped from an *RGB* face image. (c), (d), and (e) are R, G, and B channel images extracted from the *VIS* ocular image, respectively.

2. *Resolution of the unwrapped iris*: As mentioned in Section 1.3.4, Daugman's rubber sheet model unwraps the segmented iris into a rectangular entity of specific width and height. Two different normalization resolutions were tested: 64×360 and 32×180 . These resolutions were empirically chosen based on the pupillary and limbic radii observed in the *NIR* and *VIS* images.

^{||}As the iris regions in *VIS* ocular images are of very small resolution

It was observed that the R channel image unwrapped to a resolution of 64×360 , provided the best recognition performance (EER = 29%). The ROC curves obtained using the above matching considerations are shown in Figure 4.14. The low recognition performance, even with accurate segmentation on a small dataset, indicates the need for a better cross-spectral NIR-VIS iris matching algorithm.

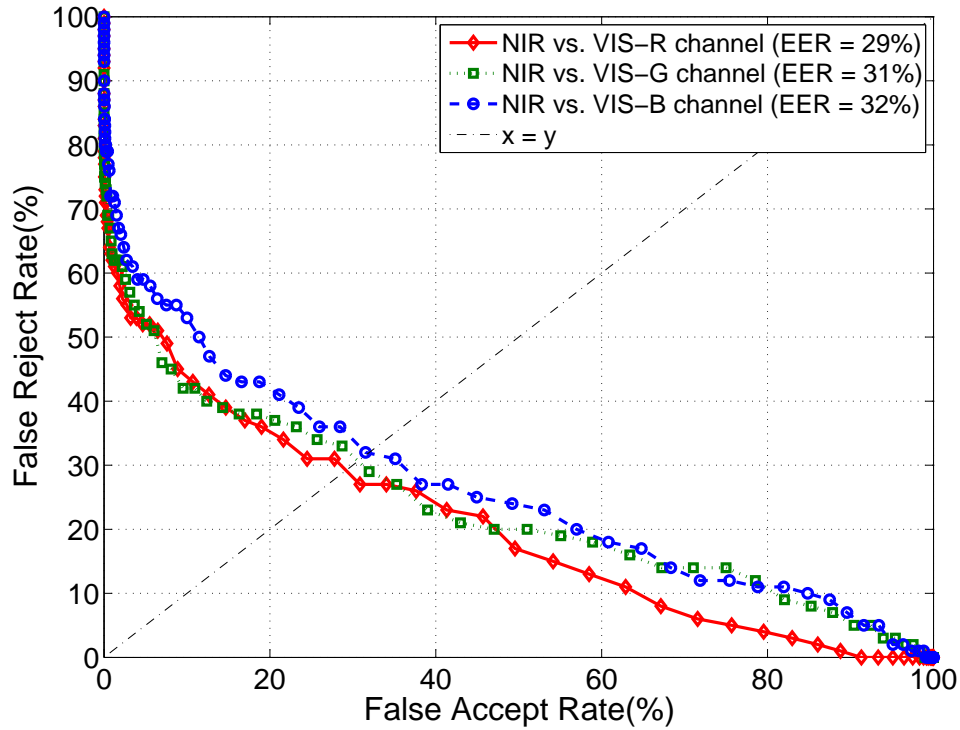


Figure 4.14: ROC curves obtained by matching iris regions extracted from R, G, and B channels of the VIS ocular images with those extracted from the NIR iris images, using Libor Masek's open source implementation. *Note that these curves correspond to matching performance obtained using a subset of images (100 NIR iris and 100 VIS ocular images).*

4.5.2 Commercial Algorithm

Owing to the poor performance of the open source implementation, a commercial iris recognition system, VeriEye [105] was used. Both *Set 1* and *Set 2* were combined to generate 1358 VIS ocular and 1358 NIR iris images from 704 subjects. Based on the performances observed in the previous experiment, only R channel images were considered. The ROC curves, along with the EERs obtained using VeriEye on the left-side images, are shown in Figure 4.15.

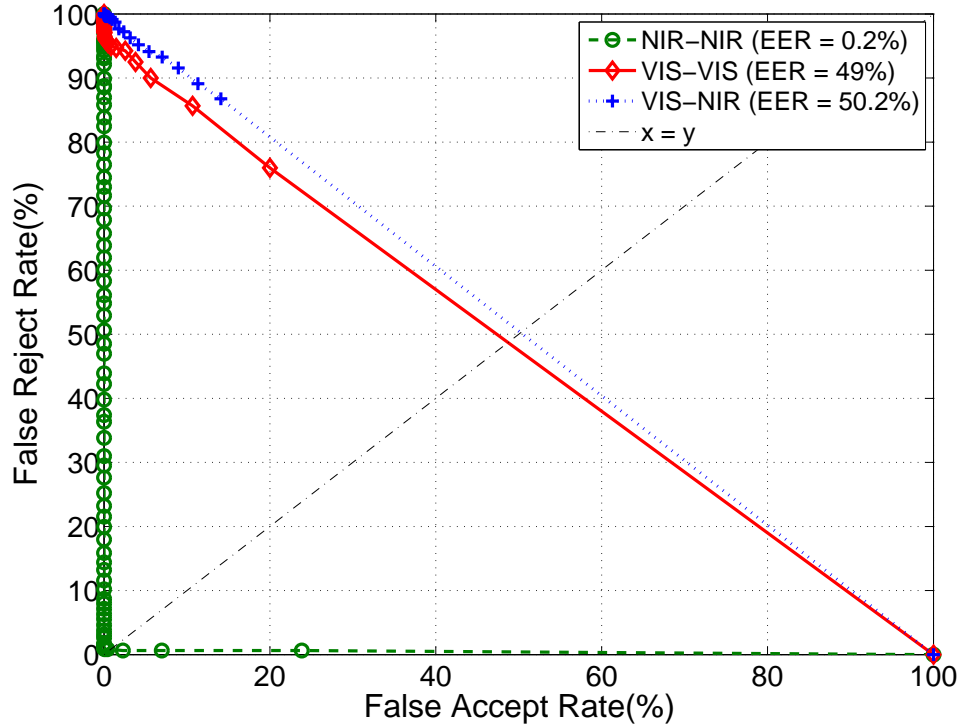


Figure 4.15: ROC curves obtained using VeriEye to match irides from left-side NIR iris images and left-side VIS ocular images.

It was observed that 74, out of the 1358 VIS ocular images, could not be processed by VeriEye. From the ROC curve, it can be observed that VeriEye provides good recognition performance only for the intra-spectral NIR-NIR iris matching. Iris matching performance using VeriEye on cross-spectral NIR-VIS and intra-spectral VIS-VIS images was very poor. The reasons for such poor performance could not be deduced as VeriEye does not provide the intermediate details of segmentation and matching.

4.6 Ocular Recognition

Based on the results presented in Sections 4.5.1 and 4.5.2, it can be observed that the iris biometric does not result in good recognition performance when cross-spectral NIR-VIS images are used. Better performance could be expected if segmentation and matching schemes are significantly improved. However, this can be a very challenging task due to the presence of multiple non-ideal factors. Ross et al. [4] and Woodard et al. [23] suggest that the ocular region can provide better recognition performance under non-ideal conditions. In

this regard, the following sections present various techniques used to perform cross-spectral NIR-VIS ocular region matching.

4.6.1 Baseline - Local Binary Patterns

From Table 4.3, it can be observed that the three most popular techniques used for ocular image matching are: Gradient Orientation Histograms [138], Local Binary Patterns [139], and Scale Invariant Feature Transform [123]. Based on an empirical evaluation on a sample dataset**, LBP was observed to provide better recognition performance on Bio-CoP images, in comparison with GOH and SIFT. Therefore, LBP was chosen as a baseline algorithm to perform ocular region recognition on the entire set of 1308 VIS and 1308 NIR images. Both the NIR and VIS images were resized to a fixed resolution of (225×169) pixels. Such resizing helps in having: (a) rough localization of the regions of interest, and (b) fixed size feature vectors. The EERs obtained using LBP on left-side images corresponding to (a) NIR-NIR, (b) VIS-VIS, and (c) VIS-NIR ocular region matching were observed to be 35%, 12%, and 50.4%, respectively. The corresponding ROC curves are provided in Figure 4.16.

From the results, it can be observed that the cross-spectral VIS-NIR ocular region recognition performance is no better than that of the iris biometric. Two main reasons for the low performance of LBP were observed to be: (a) appearance variations of the ocular regions, caused by different viewing angles of the sensors, and (b) reduced textural quality of the ocular regions in VIS images.

4.6.2 Normalized Gradient Correlation

Correlation based approaches have been observed to provide better recognition performance when compared to histogram based approaches (e.g., LBP, SIFT, GOH, etc.) on non-ideal ocular images [4] [132]. To test this observation, the Normalized Gradient Correlation (NGC) method proposed by Tzimiropoulos et al. [140] was used. NGC was initially proposed for image registration and alignment. In this work, the technique is modified to perform image matching by using the gradients within a considered region of interest. The advantages offered by NGC are that (a) it can well handle illumination variations, and (b)

**100 NIR iris and 100 VIS ocular images

Table 4.3: List of major biometrics research publications focused on the ocular region

Year	Authors	Feature extraction	Visible or NIR	Major observation
2009	Park et al. [90]	SIFT, LBP, GOH	Visible	Serves as a soft biometric
2010	Woodard et al. [23]	LBP	NIR	Aids iris recognition under non-ideal conditions
2010	Bharadwaj et al. [91]	GIST, CLBP, SIFT	Visible	Alternative to iris recognition at a distance
2010	Merkow et al. [93]	LBP with SVM	Visible	Can help in gender classification
2010	Hollingsworth et al. [95]	Human expertise	NIR	Identifying useful features in ocular images, as discerned by humans
2011	Xu et al. [131]	WLBP	Visible	Improves face recognition performance under aging
2011	Boddeti et al. [132]	PDM	NIR	Comparison of iris and ocular recognition performances
2011	Dong and Woodard [133]	Manual segmentation with LDA and SVM	Both	Analysis of eyebrow shape based features for recognition and gender classification
2011	Park et al. [24]	SIFT, LBP, GOH	Visible	Can improve face recognition when faces are masked
2012	Jillela and Ross [134]	SIFT, LBP	Visible	Can improve face recognition under plastic surgery
2012	Ross et al. [4]	SIFT, LBP, PDM	NIR	Outperforms iris recognition under highly non-ideal conditions
2012	Padole and Proenca [135]	SIFT, LBP, GOH	Visible	Analysis of various performance degradation factors
2012	Hollingsworth et al. [94]	SIFT, LBP, GOH	Both	Comparison of human and machine performances
2012	Oh et al. [136]	Variations of PCA, LDA	Visible	Performance analysis of projection-based methods
2013	Hollingsworth et al. [137]	Active Shape Models	Visible	Automatic eyebrow segmentation

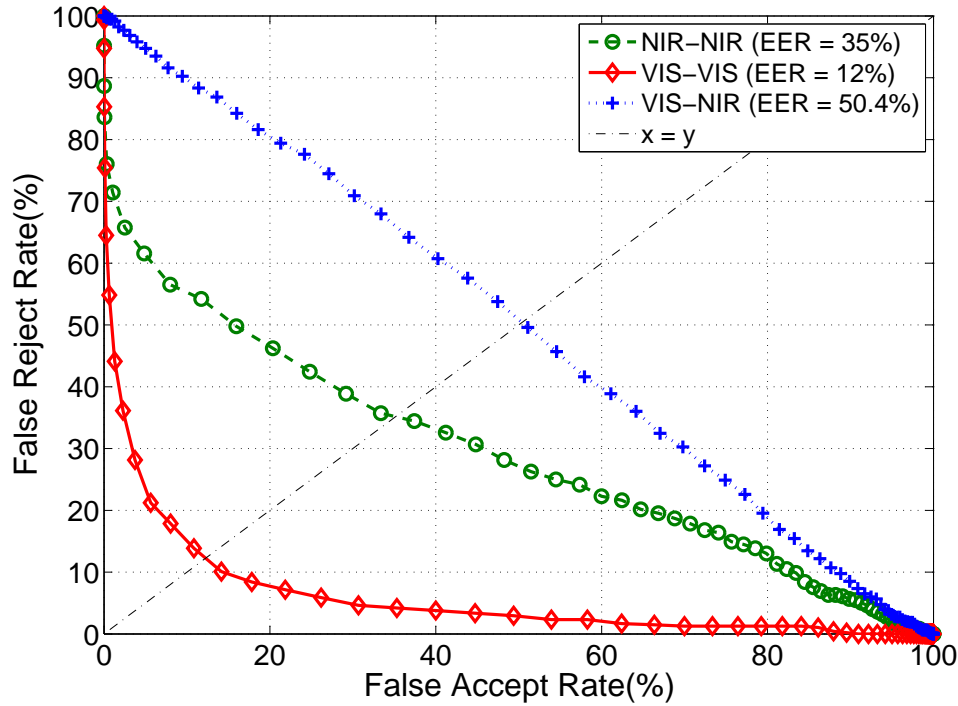


Figure 4.16: ROC curves corresponding to the ocular region matching using LBP on left-side images.

it does not require any learning.

Normalized Gradient Correlation computation between two images is similar to that of the 2D normalized cross-correlation. The only difference is that it operates on the image gradients instead of the raw pixel intensity values. Given two images I_1 and I_2 of the same resolution, the normalized cross-correlation can be computed as

$$F^{-1} \left\{ \frac{\widehat{I}_1 \widehat{I}_2^*}{|\widehat{I}_1| |\widehat{I}_2^*|} \right\} \quad (4.1)$$

where $\widehat{I}_1 = F\{I_1\}$, $\widehat{I}_2 = F\{I_2\}$, and F denotes the Fourier transform operation. On the other hand, the Normalized Gradient Correlation can be computed as

$$F^{-1} \left\{ \frac{\widehat{G}_1 \widehat{G}_2^*}{|\widehat{G}_1| |\widehat{G}_2^*|} \right\} \quad (4.2)$$

where $\widehat{G}_1 = F\{G_1\}$, $\widehat{G}_2 = F\{G_2\}$, and $G_1 = G_{1x} + G_{1y}$, $G_2 = G_{2x} + G_{2y}$. The terms G_{ix} and G_{iy} represent the gradients of the image I in x and y directions, respectively.

To perform ocular matching between two images using NGC, each image was first tessellated into 12 non-overlapping patches of equal size. The NGC value between the

corresponding pairs of patches between NIR and VIS images was computed, yielding 12 different patch scores. The value obtained by the summation of all such patch scores was used as the final match score between an image pair. The ROC curves corresponding to NGC based ocular matching are provided in Figure 4.17. The EER values corresponding to NIR-NIR, VIS-VIS, and VIS-NIR matching using left-side images were observed to be 20%, 8%, and 34%, respectively.

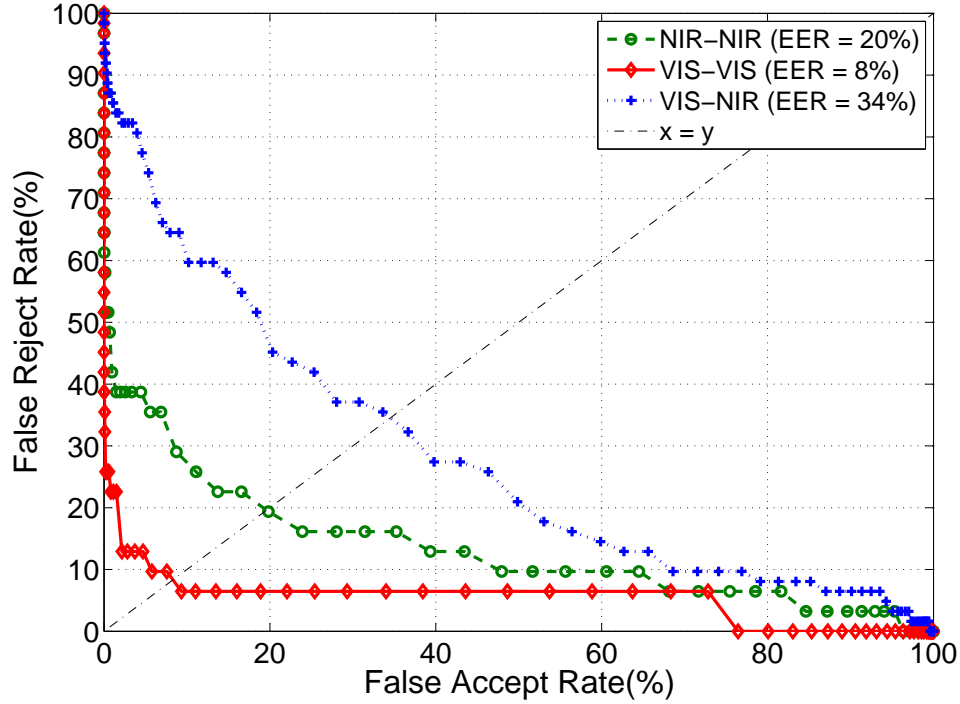


Figure 4.17: ROC curves corresponding to ocular matching using the Normalized Gradient Correlation technique on left-side images.

From the results, it can be observed that NGC provides better recognition performance for cross-spectral VIS-NIR image matching compared to (a) iris recognition, and (b) LBP based ocular region recognition. The reason for such improved performance is that the process of comparing the patches on a one-to-one basis eliminates the need for having tight correspondences between the sampling points of histogram based methods. Furthermore, using the image gradients ensures that the edge information corresponding to the shapes within the image are accounted for. It has to be noticed that the NIR-NIR ocular matching performance is also improved by NGC.

4.7 Ocular Matching using Joint Dictionary Approach

Sparse representation based approaches for image matching have been gaining significant attention in the recent past. Such approaches have been successfully applied in biometrics for face [141], iris [142] and ear [143] recognition. Sparse coding approximates a given image by a linear combination of a few atoms from a dictionary learned from a training set of images. Sparse approaches allow encoding of images into sparse vectors even under various challenging variations. Such methods have also been used in many image processing problems such as denoising [144], restoration [145], and super-resolution [146].

4.7.1 Sparse Representation Framework

The basic framework for sparse representation based approaches for pattern classification mainly depends on: (a) the dictionary formed using the training samples, and (b) the conditions used for obtaining the sparse representation of a given test sample. Consider a training dataset that contains k image samples corresponding to each of n different subjects (i.e., classes). A data matrix, A , can be obtained by concatenating all the given training images as column vectors as:

$$A = [I_{1,1}, I_{1,2}, \dots, I_{1,k}, \dots, I_{n,1}, I_{n,2}, \dots, I_{n,k}], \quad (4.3)$$

where $I_{i,j}$ denotes the i th image sample ($i = 1, 2, \dots, k$) of the j th subject ($j = 1, 2, \dots, n$), represented as a column vector. Assume that a sufficient number of training images corresponding to each class are available. A new test image, y , can then be represented as a linear representation of the data matrix entries. This process can be mathematically represented as:

$$y = \alpha_{1,1}I_{1,1} + \alpha_{1,2}I_{1,2} + \dots + \alpha_{n,k}I_{n,k}, \quad (4.4)$$

where $\alpha_{i,j}$ represents a scalar coefficient corresponding to i th image of the j th subject. The above equation can be summarized as

$$y = Ax, \quad (4.5)$$

where $x = [\alpha_{1,1}, \alpha_{1,2}, \dots, \alpha_{n,k}]$.

In an identification scenario, the identity of y can be determined by solving the following

minimization problem:

$$\hat{x} = \arg \min \|x\|_1 \text{ subject to } Ax = y. \quad (4.6)$$

The coefficient vector, \hat{x} , typically contains non-zero entries that correspond to the identity of the test sample, and zeros everywhere else. The generic structure of \hat{x} can therefore be given as

$$\hat{x} = [0, \dots, 0, \alpha_{p,1}, \alpha_{p,2}, \dots, \alpha_{p,n}, 0, \dots, 0], \quad (4.7)$$

where p corresponds to the true identity of y . In the presence of noise, a stable solution can be determined by rewriting Equation (4.6) as:

$$\hat{x} = \arg \min \|x\|_1 \text{ subject to } \|Ax - y\|_2 \leq \epsilon, \quad (4.8)$$

where ϵ represents a desired threshold.

The above framework was modified by Guo et al. [147] to perform *verification* using face images. Given a pair of face images I_p and I_q , their sparse representation vectors, \hat{x}_p and \hat{x}_q , are first computed using Equation (4.8). The Euclidean distance between \hat{x}_p and \hat{x}_q is then used to determine the similarity between the two images.

It has to be noted that the data matrix A is typically referred as an *overcomplete dictionary* whose base elements are the training images themselves. This leads to a large dimensionality of A , resulting in expensive computations. A large number of algorithms have been proposed for learning a compact dictionary while ensuring sparsity [148] [149]. One such method [150] to determine a compact dictionary D , from a given training data matrix A involves the following equation:

$$D = \arg \min_{D,Z} \|A - DZ\|_2 + \lambda \|Z\|_1, \quad (4.9)$$

where Z and λ represent the sparse coefficient matrix and the regularization parameter, respectively,

4.7.2 Joint Dictionary Approach

It has to be noted that the above described sparse representation framework may not be directly applicable to the current VIS-NIR ocular image matching problem. A vast majority of the existing approaches generate a single dictionary using the training images. To

perform VIS-NIR ocular image matching, however, two different dictionaries are required. This is because the sparse representations for even the genuine pairs of VIS and NIR images can be significantly different due to variations in image acquisition.

Consider two different dictionaries D_{NIR} , and D_{VIS} , generated using NIR only, and VIS only training images, respectively. Let I_{NIR} and I_{VIS} represent a pair of NIR and VIS ocular images that have to be matched. The sparse representation vector of I_{VIS} , represented by \hat{x}_{VIS} , can be computed as:

$$\hat{x}_{VIS} = \arg \min \|x_{VIS}\|_1 \text{ subject to } \|D_{VIS}x_{VIS} - I_{VIS}\|_2 \leq \epsilon_{VIS}. \quad (4.10)$$

Similarly, the corresponding sparse representation vector of I_{NIR} , denoted by \hat{x}_{NIR} , can be computed as:

$$\hat{x}_{NIR} = \arg \min \|x_{NIR}\|_1 \text{ subject to } \|D_{NIR}x_{NIR} - I_{NIR}\|_2 \leq \epsilon_{NIR}. \quad (4.11)$$

The similarity between \hat{x}_{NIR} and \hat{x}_{VIS} can not be directly used as a measure of the similarity between the images due to the differences in D_{NIR} and D_{VIS} . If the relation between NIR and VIS images could be modeled, D_{NIR} and D_{VIS} could then be related to each other. However, such a modeling is very difficult due to a multitude of factors that cause variations within NIR and VIS images. This problem can be mitigated by combining the dictionaries D_{NIR} and D_{VIS} by a joint dictionary training approach. Such an approach ensures that \hat{x}_{NIR} and \hat{x}_{VIS} have similar non-zero coefficients if I_{NIR} and I_{VIS} correspond to the same subject.

4.7.3 Dictionary Learning and Matching

Consider a data matrix, A_{VIS} , generated from a set of VIS images, using Equation (4.3):

$$A_{VIS} = [I_{1,1}^{VIS}, I_{1,2}^{VIS}, \dots, I_{1,k}^{VIS}, \dots, I_{n,1}^{VIS}, I_{n,2}^{VIS}, \dots, I_{n,k}^{VIS}]. \quad (4.12)$$

Let the corresponding NIR images be used to generate A_{NIR} :

$$A_{NIR} = [I_{1,1}^{NIR}, I_{1,2}^{NIR}, \dots, I_{1,k}^{NIR}, \dots, I_{n,1}^{NIR}, I_{n,2}^{NIR}, \dots, I_{n,k}^{NIR}]. \quad (4.13)$$

A number of approaches have been proposed for effective dictionary learning [148]. The formulation used in this work is inspired by the joint dictionary learning approach proposed

by Yang et al. [151] for image super-resolution. The *independent* compact dictionaries for VIS and NIR images, D_{VIS} and D_{NIR} , can be determined by:

$$D_{VIS} = \arg \min_{D_{VIS}, Z} \|A_{VIS} - D_{VIS} * Z_{VIS}\|_2 + \lambda_{VIS} \|Z_{VIS}\|_1, \quad (4.14)$$

and

$$D_{NIR} = \arg \min_{D_{NIR}, Z} \|A_{NIR} - D_{NIR} * Z_{VIS}\|_2 + \lambda_{NIR} \|Z_{NIR}\|_1, \quad (4.15)$$

where Z and λ represent the sparse coefficient matrix and regularization parameter, respectively, for the considered set of VIS or NIR test images.

The goal here is to learn a *joint* dictionary such that the sparse representation of an NIR test image will be similar to that of its corresponding VIS image of the same subject. Therefore, Equations (4.14) and (4.15) can be combined as:

$$\arg \min_{D_{NIR}, D_{VIS}, Z} \|A_{VIS} - D_{VIS} * Z\|_2 + \|A_{NIR} - D_{NIR} * Z\|_2 + \lambda \|Z\|_1. \quad (4.16)$$

The above equation could be rewritten as:

$$\arg \min_{D_{NIR}, D_{VIS}, Z} \|A_{joint} - D_{joint} * Z\|_2 + \lambda \|Z\|_1, \quad (4.17)$$

where $A_{joint} = \begin{bmatrix} A_{VIS} \\ A_{NIR} \end{bmatrix}$ and $D_{joint} = \begin{bmatrix} D_{VIS} \\ D_{NIR} \end{bmatrix}$.

Efficiently solving the above formulation using numerical methods is a challenge by itself. To this end, multiple solutions have been proposed in the machine learning domain [148]. In this work, the approach suggested by [150] is used. Equation (4.17) is considered to be non-convex in both D and Z collectively, but is convex in one of them if the other is fixed. Therefore, the optimization is performed in an alternate manner over D_{joint} and Z . The optimization algorithm is outlined in Algorithm 1. MATLAB packages provided by [150] and [151] were used for solving the algorithm. A variation of the joint dictionary approach has been used by Shekhar et al. [152]. However, such techniques have been used for identification and not for verification.

The proposed VIS-NIR ocular image matching technique is outlined in Algorithm 2. 30% of the database was used for training and the remaining 70% was used for testing (disjoint subjects). This results in considering 407 and 951 images for training and testing, respectively. The obtained match scores are used as similarity measures between the given images. The ROC curves obtained using the proposed joint dictionary based sparse

Algorithm 1: Joint Dictionary Learning*Step 1*Use a Gaussian random matrix to initialize D_{joint} *Step 2*With D_{joint} fixed, update Z by solving the following formulation:

$$Z = \arg \min_Z \|A_{joint} - D_{joint} * Z\|_2 + \lambda \|Z\|_1$$

*Step 3*With Z fixed, update D_{joint} by:

$$D_{joint} = \arg \min_{D_{joint}} \|A_{joint} - D_{joint} * Z\|_2 \text{ such that } \|D_{joint}\|_2 \leq 1$$

Step 4

Iterate between steps 2 and 3 until convergence.

Final Output: D_{joint} **Algorithm 2: Proposed VIS-NIR Ocular Image Matching Approach***Training*

1. Input: VIS and corresponding NIR training image pairs
2. Obtain $D_{joint} = \begin{bmatrix} D_{VIS} \\ D_{NIR} \end{bmatrix}$ using Algorithm 1.

Testing

1. Input: Given VIS and NIR test images, I_{VIS} and I_{NIR}
2. Compute the sparse representation vectors \hat{x}_{VIS} and \hat{x}_{NIR} (use Equations (4.10) & (4.11))
3. Compute the Euclidean distance, d , between \hat{x}_{VIS} and \hat{x}_{NIR}
4. Determine a vector K whose entries satisfy the condition:
 $\{\hat{x}_{VIS}(k) > 0 \text{ and } \hat{x}_{NIR}(k) > 0\}$ or $\{\hat{x}_{VIS}(k) < 0 \text{ and } \hat{x}_{NIR}(k) < 0\}$
5. Match score between I_{VIS} and I_{NIR} is considered as $d/\text{size}(K)$

Final Output: $d/\text{size}(K)$

representation approach on the left-side images are provided in Figure 4.18. The EERs obtained using all the ocular matching techniques considered in this work are listed in Table 4.4. From the results it can be noticed that the proposed joint dictionary based sparse representation approach improves the recognition performance in all the three matching scenarios (i.e., NIR-NIR, VIS-VIS, and VIS-NIR).

Table 4.4: **Equal Error Rates obtained using left-side images of the considered BioCoP database.**

	NIR-NIR	VIS-VIS	VIS-NIR
Iris Recognition - VeriEye	0.2%	49%	50.2%
Ocular Recognition - Local Binary Patterns (LBP)	35%	12%	50.4%
Ocular Recognition - Normalized Gradient Correlation (NGC)	20%	8%	34%
Ocular Recognition - Joint Dictionary based Sparse Representation (JDSR) approach	14%	7%	26%

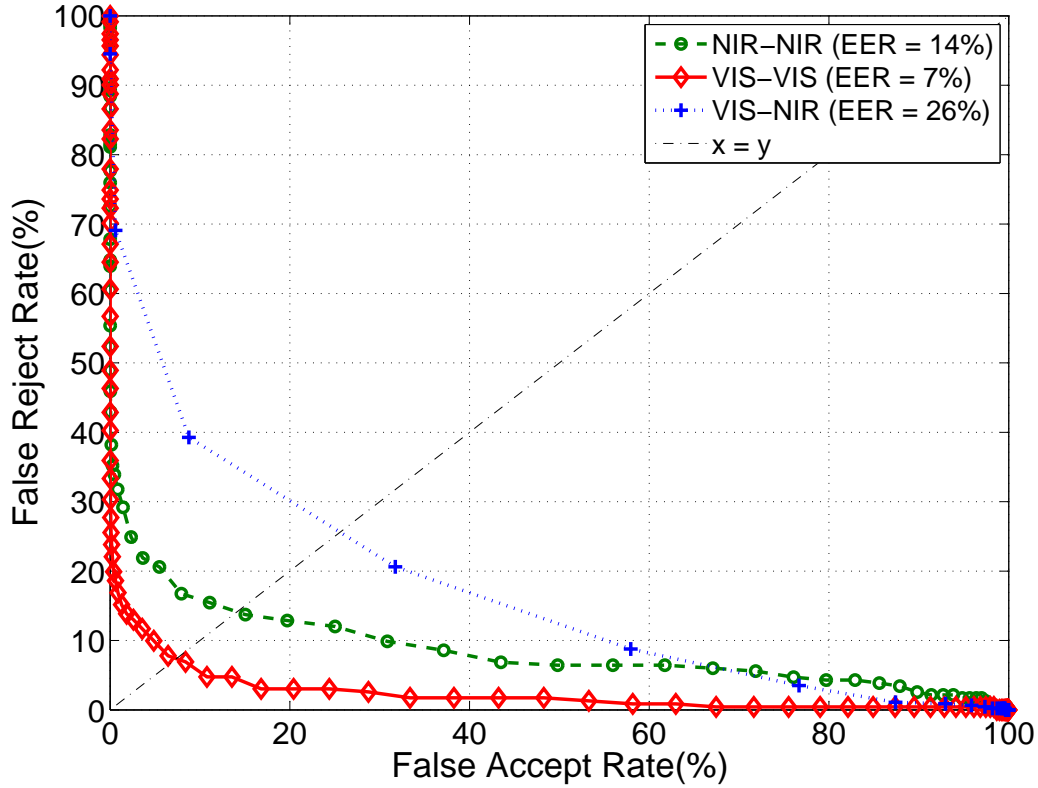


Figure 4.18: ROC curves corresponding to ocular matching using the joint dictionary based sparse representation approach on left-side images.

4.8 Computational Details

The value of λ was set to 0.09 in this work. This was done based on observing the recognition accuracies corresponding to different values of λ . The value corresponding to minimal EER was chosen for experiments. It has to be noted that other methods exist to choose the value of λ (e.g., based on the objective function, based on the sparsity of the vector, etc.). However, since the focus of this work is on recognition performance, λ is chosen empirically. The EERs obtained for various values of λ are listed in Table 4.5. The time required for generating the sparse representation of a given image was observed to be 0.8 seconds using an Intel Core i7 processor with a 3.4 GHz processor and 8 GB RAM.

4.9 Score-level fusion

To study the impact of score-level fusion, match scores obtained by all the three considered ocular matching techniques were combined using a simple sum rule. LBP and

Table 4.5: **Equal Error Rates corresponding to different values of λ .**

λ value	EERs for VIS-NIR matching
0.01	36.5%
0.05	33.1%
0.09	26.0%
0.15	36.2%

NGC scores that correspond only to those image pairs used in the testing phase of the joint dictionary based sparse approach were fused. The weights for fusion were determined empirically with an objective of minimizing the EER. The ROC curves obtained using the individual techniques and by the score-level fusion for VIS-NIR ocular matching are shown in Figure 4.19. From the results, it can be observed that score-level fusion enhances the recognition performance only by a small margin (EER = 23%). This is because of the low recognition performances offered by LBP and NGC techniques.

4.10 Summary

The problem of matching ocular regions in *RGB* face images and *NIR* iris images is studied. The variations in modalities, wavelengths, resolutions, and sensors render this problem very challenging. A sparse representation based approach which generates a joint dictionary from corresponding pairs of ocular regions in NIR and VIS images is proposed. The proposed technique is observed to outperform some of the well known ocular matching techniques. Additionally, this work highlights the potential of ocular region in non-ideal conditions when iris information may not be reliable. Future work would include investigating the robustness of the proposed approach when accurate localization of the ocular regions is not possible. Also, the current work does not account for geometric deformations that can occur in the ocular region. A robust ocular matching model that can simultaneously handle photometric and geometric variations has to be developed.

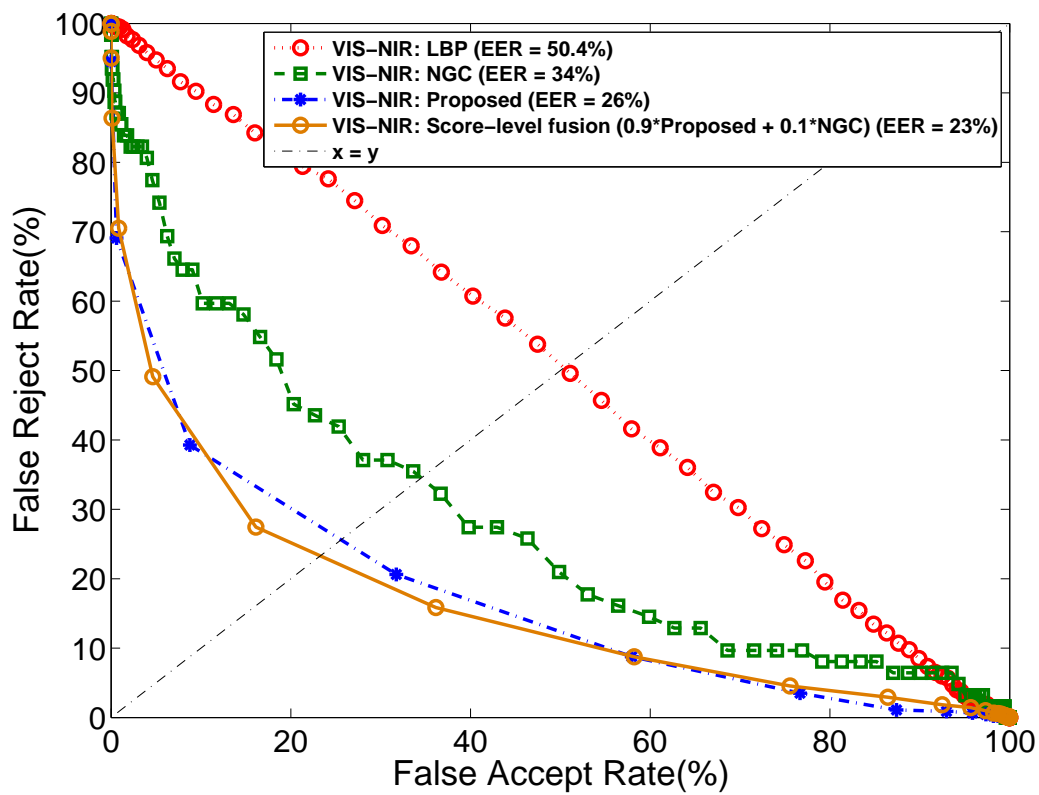


Figure 4.19: ROC curves for the cross-spectral ocular matching using (a) LBP, (b) NGC, (c) proposed joint dictionary based sparse representation approach, and (d) weighted score-level fusion of all the considered techniques.

Chapter 5

Summary and Conclusions

Iris is considered to be one of the most accurate and reliable biometric traits. However, high iris recognition performances are typically observed under constrained conditions. Under unconstrained conditions, the surrounding ocular region information can be used to improve the overall recognition performance. Ocular recognition research is still in its incipient stages. Much like iris, ocular region performance can be negatively impacted by several non-ideal factors. This thesis focused on improving iris and ocular region recognition performances under non-ideal conditions.

5.1 Research Contributions

Chapter 2 considers the problem of low-resolution iris recognition and provides the following research contributions:

- The impact of systematically lowering the image resolution on iris recognition performance was studied.
- An image-level fusion technique based on Principal Components Transform was proposed. It was demonstrated that the proposed technique can successfully fuse information and improve recognition performance of low-resolution iris video frames.
- A comparison of image-level and score-level fusion in low-resolution iris imagery is provided. Score-level fusion is observed to perform better than image-level fusion in the considered scenario.

In Chapter 3, the impact of plastic surgeries on face recognition was studied and the following contributions were made:

- A score-level fusion scheme was proposed that combines information from both face and ocular regions. The proposed technique improved the overall recognition performance, thereby highlighting the usage of ocular region in challenging conditions.
- The effect of individual plastic surgeries on face recognition was studied. It was demonstrated that the proposed fusion scheme improves the overall recognition performance for all considered individual surgeries.

Chapter 4 presents the problem of matching ocular regions extracted from face and iris images. The research contributions offered by this work can be listed as follows:

- The first study on ocular matching using *RGB* face and *NIR* iris images was presented.
- Ocular region was demonstrated to outperform iris recognition in the considered database. This suggests the importance of ocular region in improving the recognition performance under non-ideal conditions.
- An ocular image matching technique based on sparse representation approach was proposed. The proposed technique is shown to provide better recognition performance than the existing ocular recognition techniques.

5.2 Directions for Future Research

Based on the overall experience gained from all the chapters listed in this thesis, the following directions could be used for future research* in ocular recognition:

- *Iris unwrapping*: In Chapter 2, all the iris regions extracted from images of varying resolutions are unwrapped to rectangular entities of fixed dimension. Fixed resolution sampling and unwrapping can induce significant levels of noise when low-resolution iris images are used. This noise can be reduced by maintaining the dimensions of the unwrapped iris proportional to the difference between limbic and pupillary radii. However, such a setup does not allow for easy matching between unwrapped images extracted from varying resolution iris images. Further research is required to investigate and mitigate this problem.

*Please note that the potential improvements for the individual problems considered in this thesis are provided at the end of their respective chapters.

- *Iris segmentation*: As shown in Chapter 4, iris segmentation in images corresponding to varying illumination, resolutions, and wavelengths is still a very challenging problem. Although a significant number of non-ideal iris segmentation approaches have been proposed in the literature, their computational cost and processing times remain high. In this regard, future research should be aimed towards developing robust iris recognition techniques that are segmentation independent.
- *Adaptive ocular trait selection*: Ocular region has been proven to aid, or even outperform iris recognition in non-ideal scenarios. However, such observations have been made only when iris information cannot be reliably acquired. In this regard, an adaptive ocular trait selection scheme that automatically decides the region of interest (e.g., iris, ocular region, and conjunctiva) can be of significant help in maintaining a reliable performance.
- *Cross-spectral matching*: Performing ocular matching using images acquired under varying spectra of light is still an unsolved problem. Modeling the photometric and geometric variations between the images, rather than invoking simple photometric correction techniques, is an attractive area of research.

APPENDICES

Appendix A

Methods for Iris Segmentation

A.1 Classical Iris Segmentation Algorithms

A.1.1 Integro-Differential Operator

The technique proposed by John Daugman [62] in the early 1990s is considered to be the pioneering work in the field of automated iris recognition. A vast majority of commercial iris recognition systems worldwide, employ Daugman's approach. In this approach, iris segmentation is carried out using an integro-differential operator. The segmentation process is performed by approximating the iris boundaries as perfect circles.

Given an iris image $I(x, y)$, it is first convolved with an image smoothing function (e.g., a Gaussian filter). This process of smoothening the image helps in (a) attenuating the effect of noise (e.g., sensor noise) in the image, and (b) eliminating undesired weak edges (e.g., boundaries within the iris), while retaining the desired strong edges (e.g., iris boundaries, eyelid boundaries, etc.). An integro-differential operator is then used to search for the maximum value of a normalized integral along circular contours of varying radii and center coordinates. The search process over the image domain (x, y) using an integro-differential operator can be mathematically expressed as:

$$\max(r, x_0, y_0) \left| G_\sigma(r) * \frac{\partial}{\partial r} \oint_{r, x_0, y_0} \frac{I(x, y)}{2\pi r} ds \right|, \quad (\text{A.1})$$

where

$$G_\sigma(r) = \frac{1}{\sqrt{2\pi}\sigma} \exp^{-\left(\frac{(r-r_0)^2}{2\sigma^2}\right)} \quad (\text{A.2})$$

represents the radial Gaussian with a center r_0 and standard deviation (scale) σ , which is used for image smoothing. The symbol $*$ denotes the convolution operation, and r represents the radius of the circular arc ds , centered at the location (x_0, y_0) . The division

by a factor of $2\pi r$ normalizes the circular integral with respect to its perimeter. In other words, the integro-differential operator behaves as a circular edge detector, that searches iteratively for the maximum response of a contour path defined by the parameters (x_0, y_0, r) . Depending on the values of the radii considered, the optimal parameters of the integro-differential operator are treated as either the pupillary or limbus boundaries. Figure A.1 illustrates the search process using an integro-differential operator.

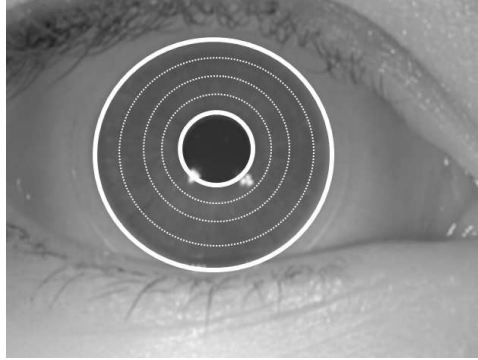


Figure A.1: Iris segmentation using the integro-differential operator.

The value of σ , which controls the amount of blurring of the iris image, can be varied when searching for the pupillary and the limbus boundaries. As the pixel intensity variation across the pupillary boundary is more pronounced, the σ value can be set for a coarse scale of convolution. On the other hand, when the search process is carried out for the limbus boundary, the σ value is set for a finer convolution scale. This is due to the nominal variation of the pixel intensities across the limbus boundary.

In an iris image acquired under near-ideal conditions from a cooperative subject, both the iris boundaries can be easily detected using the integro-differential operator. However, in an image acquired under non-ideal conditions, the limbus boundary may not be completely circular due to the occlusions caused by the eyelids. Therefore, when searching for the limbus boundary, the angular arc of integration, ds , is often restricted to the left and right quadrants (i.e., near the vertical edges of the iris). When searching for the pupillary boundary, this arc can be extended over a wider range, as the eyelid occlusions are relatively small. Once both the iris boundaries are detected, the boundaries of the eyelids can be detected by changing the integration path of the operator from circular to arcuate. Figure A.2 shows the output of the integro-differential operator when used to detect both

the iris boundaries and the eyebrows.

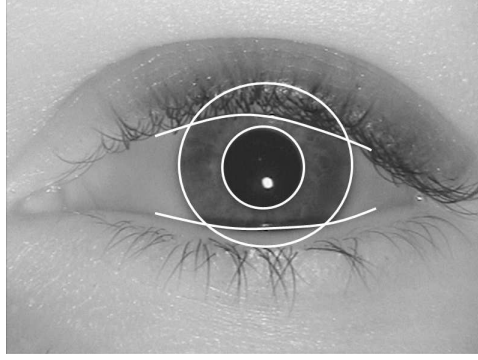


Figure A.2: Output obtained by applying an integro-differential operator to detect both the iris and eyebrow boundaries.

Despite successfully determining the iris boundaries using an integro-differential operator, the localized iris may be further occluded by other noisy regions such as eyelashes, shadows, or specular reflections. Therefore, a noise mask* that records the locations of these undesired occlusions is correspondingly generated. This mask is later used during the matching stage to mitigate the effect of noisy pixels.

A.1.2 Hough Transform

Another widely used classical iris segmentation algorithm was proposed by Wildes et. al. [51, 153]. To detect each iris boundary, the algorithm relies on the Hough transform: a histogram based model fitting approach. First, an edge map of the input image is generated using a gradient-based edge detector. A voting procedure is then applied on the highlighted edge map, to determine the parameter values for a contour that best fits a circle.

Given an iris image $I(x, y)$, the edge map of an input image can be highlighted by thresholding the magnitude of the image intensity gradient. This operation can be mathematically expressed as:

$$|\nabla G(x, y) * I(x, y)| \geq th, \quad (\text{A.3})$$

where $\nabla \equiv (\partial/\partial x, \partial/\partial y)$, and th denotes an empirically chosen intensity threshold. $G(x, y)$ represents a two dimensional Gaussian with center, (x_0, y_0) , and standard deviation, σ , used

*The process of generating a noise mask, and the subsequent schemes for iris normalization and matching are very similar in a majority of iris recognition algorithms. However, as this work focuses only on iris segmentation, these details are not discussed. The reader is directed to the original publication by Daugman [62] for further information.

for smoothing the image. The purpose of image smoothing is to (a) select the spatial scale of edges under consideration, and (b) reduce the effect of noise on the thresholding process. The mathematical expression for the Gaussian is:

$$G(x, y) = \frac{1}{2\pi\sigma^2} \exp^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\sigma^2}}. \quad (\text{A.4})$$

The image thresholding operation yields an approximate edge map consisting of the iris boundary (along with other prominent edges). In most cases, such an output would consist of non-continuous, non-circular contours. Therefore, the edge map is thinned using a morphological operation, and a voting procedure is used to determine the parameters of the iris boundaries. Hough transform [154], a standard machine vision technique for fitting simple contour models to images, is typically used during the voting process.

When searching for the limbus boundary contour, the image intensity derivatives corresponding to vertical edges are weighted more during the voting process. This directional selectivity grants preference to the left and right portions of the limbus boundary over its upper and lower portions. Therefore, even if the upper and lower portions of the limbus boundary are occluded by eyelids, the left and right portions remain clearly visible and oriented (assuming the head is in an upright position).

Consider a set of edge points (x_j, y_j) , $j = \{1, 2, \dots, n\}$, obtained by the image thresholding operation. The goal is to determine if a subset of these points are associated with the contour of a circle. A circle can be parameterized as (x_c, y_c, r) where (x_c, y_c) denotes its center and r denotes its radius. The Hough transform detects circular contours in the edge image by defining an accumulator array, H , whose entries $H(x_c, y_c, r)$ are computed as follows:

$$H(x_c, y_c, r) = \sum_{j=1}^n h(x_j, y_j, x_c, y_c, r) \quad (\text{A.5})$$

where

$$h(x_j, y_j, x_c, y_c, r) = \begin{cases} 1, & \text{if } g(x_j, y_j, x_c, y_c, r) = 0 \\ 0, & \text{otherwise.} \end{cases} \quad (\text{A.6})$$

and

$$g(x_j, y_j, x_c, y_c, r) = (x_j - x_c)^2 + (y_j - y_c)^2 - r^2. \quad (\text{A.7})$$

For each edge point (x_j, y_j) , $g(x_j, y_j, x_c, y_c, r)$ is set to 0 if the parameter triplet (x_c, y_c, r) represents a circle through that point. The parameter triplet that maximizes H is consid-

ered to be a reasonable choice to represent the contour of interest. The maximizing parameter triplet is determined by first building $H(x_c, y_c, r)$ as an array (indexed by discretized values for x_c , y_c , and r), and then scanning for the triplet that corresponds to the largest value in the array.

The same process (consisting of thresholding and voting) is used to determine the pupillary boundary, but with the following minor modifications:

1. The image is filtered with a gradient-based edge detector that is *not* directionally tuned. This is due to the fact that the pupillary boundary is less prone to occlusion from the eyelids.
2. The permissible parameter values (x_c, y_c, r) are constrained to lie within the circle that describes the limbus boundary.

Once both the iris boundaries are detected, it is necessary to determine the locations of the upper and lower eyelids that may occlude the iris. To perform this operation, a gradient-based edge detector that is tuned to favor horizontal edges is used. This is based on the fact that the contour of the upper and lower eyelids within the limbus boundary would be nearly horizontal (under the assumption that the subject's head is in an upright position). The upper and lower eyelids are modeled as two separate parabolic arcs of the form $x(t) = a_x t^2 + b_x t + c_x$, and $y(t) = a_y t^2 + b_y t + c_y$ with $0 \leq t \leq 1$. The parameters a_x, b_x, c_x, a_y, b_y , and c_y , are once again determined using the same histogram-based model-fitting approach.

A.2 Other Prominent Approaches

Both the aforementioned techniques approximate an iris boundary with a circle or an ellipse. However, such an approximation may not be always suitable. Iris recognition performance can be improved by determining the precise boundaries of the iris, rather than their approximations. Determining the precise boundaries of the iris reduces the noise from occlusions, especially those caused by eyelashes. Figure A.3 shows the difference between an approximated and a precisely determined iris boundary.

Methods that use curve evolution processes (e.g., Geodesic Active Contours, level sets, etc.) could be extremely beneficial in determining the precise boundaries of an iris. In this

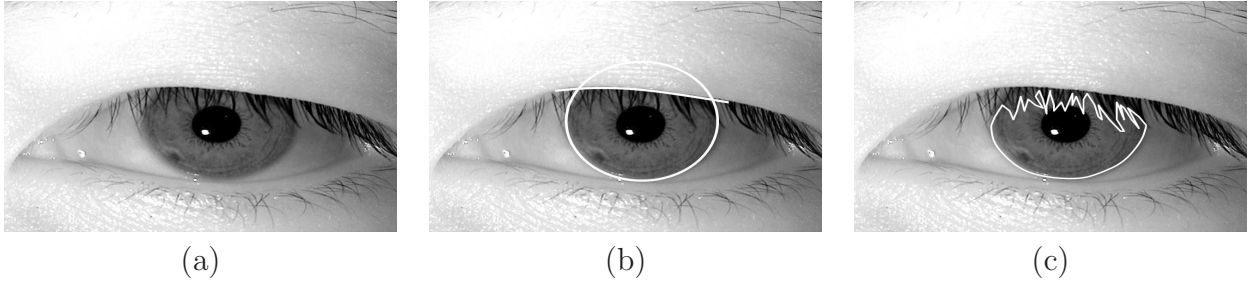


Figure A.3: (a) An eye image in which the limbus boundary is occluded by the eyelashes. (b) Approximate limbus boundary obtained using Daugman’s integro-differential operator. (c) Precise limbus boundary obtained by a curve evolution technique. Notice that the precise segmentation helps in avoiding the noise caused by eyelash occlusion.

section, some such techniques are discussed.

A.2.1 Geodesic Active Contours

This approach, proposed by Shah and Ross [67], is based on the relation between active contours and the computation of geodesics (minimal length curves). The strategy is to evolve an arbitrarily initialized curve from within the iris under the influence of geometric properties of the iris boundary. GACs combine the energy minimization approach of the classical “snakes” and the geometric active contours based on curve evolution.

Let $\gamma(t)$ be the curve, that has to gravitate toward the outer boundary of the iris, at a particular time t . The time t corresponds to the iteration number. Let ψ be a function measuring the signed distance from the curve $\gamma(t)$. That is, $\psi(x, y) = \text{distance of point } (x, y) \text{ to the curve } \gamma(t)$.

$$\psi(x, y) = \begin{cases} 0 & \text{if } (x, y) \text{ is on the curve;} \\ < 0 & \text{if } (x, y) \text{ is inside the curve;} \\ > 0 & \text{if } (x, y) \text{ is outside the curve.} \end{cases} \quad (\text{A.8})$$

Here, ψ is of the same dimension as that of the eye image $I(x, y)$. The curve $\gamma(t)$ is called the level set of the function ψ . Level sets are the set of all points in ψ where ψ is some constant. Thus $\psi = 0$ is the zeroth level set, $\psi = 1$ is the first level set, and so on. ψ is the implicit representation of the curve $\gamma(t)$ and is called the embedding function since it embeds the evolution of $\gamma(t)$. The embedding function evolves under the influence of image gradients and the region’s characteristics so that the curve $\gamma(t)$ approaches the desired boundary of the iris. The initial curve $\gamma(t)$ is assumed to be a circle of radius r just beyond the pupillary boundary. Let the curve $\gamma(t)$ be the zeroth-level set of the embedding

function. This implies that

$$\frac{d\psi}{dt} = 0.$$

By the chain rule,

$$\frac{d\psi}{dt} = \frac{\partial\psi}{\partial x} \frac{dx}{dt} + \frac{\partial\psi}{\partial y} \frac{dy}{dt} + \frac{\partial\psi}{\partial t},$$

i.e.

$$\frac{\partial\psi}{\partial t} = -\nabla\psi \cdot \gamma'(t).$$

Splitting the $\gamma'(t)$ in the normal ($N(t)$) and tangential ($T(t)$) directions,

$$\frac{\partial\psi}{\partial t} = -\nabla\psi \cdot (v_N N(t) + v_T T(t)).$$

Now, since $\nabla\psi$ is perpendicular to the tangent to $\gamma(t)$,

$$\frac{\partial\psi}{\partial t} = -\nabla\psi \cdot (v_N N(t)). \quad (\text{A.9})$$

The normal component is given by

$$N = \frac{\nabla\psi}{\|\nabla\psi\|}.$$

Substituting this in Equation (A.9),

$$\frac{\partial\psi}{\partial t} = -v_N \|\nabla\psi\|.$$

Let v_N be a function of the curvature of the curve κ , stopping function K (to stop the evolution of the curve) and the inflation force c (to evolve the curve in the outward direction) such that,

$$\frac{\partial\psi}{\partial t} = -(\text{div}(K \frac{\nabla\psi}{\|\nabla\psi\|}) + cK) \|\nabla\psi\|.$$

Thus, the evolution equation for ψ_t^\dagger such that $\gamma(t)$ remains the zeroth level set is given by

$$\psi_t = -K(c + \epsilon\kappa) \|\nabla\psi\| + \nabla\psi \cdot \nabla K, \quad (\text{A.10})$$

where, K , the stopping term for the evolution, is an image dependant force and is used to decelerate the evolution near the boundary; c is the velocity of the evolution; ϵ indicates the degree of smoothness of the level sets; and κ is the curvature of the level sets computed as

$$\kappa = -\frac{\psi_{xx}\psi_y^2 - 2\psi_x\psi_y\psi_{xy} + \psi_{yy}\psi_x^2}{(\psi_x^2 + \psi_y^2)^{\frac{3}{2}}}.$$

[†]The subscript t denotes the iteration number

Here, ψ_x is the gradient of the image in the x direction; ψ_y is the gradient in the y direction; ψ_{xx} is the 2^{nd} order gradient in the x direction; ψ_{yy} is the 2^{nd} order gradient in the y direction; and ψ_{xy} is the 2^{nd} order gradient, first in the x direction and then in the y direction. Equation (A.10) is the level set representation of the geodesic active contour model. This means that the level-set C of ψ is evolving according to

$$C_t = K(c + \epsilon\kappa)\vec{N} - (\nabla K \cdot \vec{N})\vec{N} \quad (\text{A.11})$$

where \vec{N} is the normal to the curve. The term $\kappa\vec{N}$ provides the smoothing constraints on the level sets by reducing their total curvature. The term $c\vec{N}$ acts like a balloon force and it pushes the curve outward towards the object boundary. The goal of the stopping function is to slow down the evolution when it reaches the boundary. However, the evolution of the curve will terminate only when $K = 0$, i.e., near an ideal edge. In most images, the gradient values will be different along the edge, thus requiring the use of different K values. In order to circumvent this issue, the third geodesic term $((\nabla K \cdot \vec{N}))$ is necessary so that the curve is attracted toward the boundary (∇K points toward the middle of the boundary). This term makes it possible to terminate the evolution process even if (a) the stopping function has different values along the edges, and (b) gaps are present in the stopping function.

The stopping term used for the evolution of level sets is given by

$$K(x, y) = \frac{1}{1 + \left(\frac{\|\nabla(G(x, y) * I(x, y))\|}{k} \right)^\alpha} \quad (\text{A.12})$$

where $I(x, y)$ is the image to be segmented, $G(x, y)$ is a Gaussian filter, and k and α are constants. As can be seen, $K(x, y)$ is not a function of t .

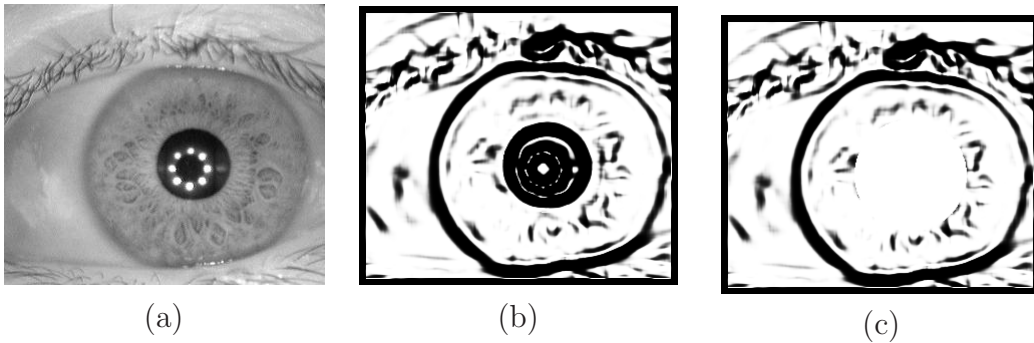


Figure A.4: Stopping function for the geodesic active contours. (a) Original iris image, (b) stopping function K , and (c) modified stopping function K' .

Consider an iris image to be segmented as shown in Figure A.4 (a). The stopping function K obtained from this image is shown in Figure A.4 (b) (for $k = 2.8$ and $\alpha = 8$). Assuming that the inner iris boundary (i.e., the pupillary boundary) has already been detected, the stopping function K is modified by deleting the circular edges corresponding to the pupillary boundary, resulting in a new stopping function K' . This ensures that the evolving level set is not terminated by the edges of the pupillary boundary (Figure A.4 (c)).

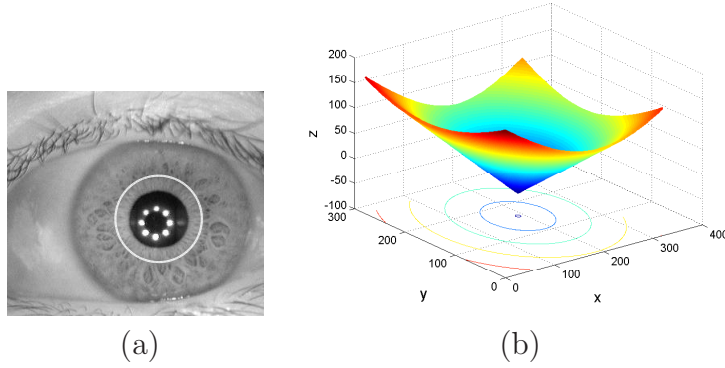


Figure A.5: Contour initialization for iris segmentation using GAC. (a) Zeroth level set (initial contour), (b) mesh plot denoting the signed distance function ψ .

A contour is first initialized near the pupil (Figure A.5 (a)). The embedding function ψ is initialized as a signed distance function to $\gamma(t = 0)$ which looks like a cone (Figure A.5 (b)). Discretizing equation A.10 leads to the following equation:

$$\frac{\psi_{i,j}^{t+1} - \psi_{i,j}^t}{\Delta t} = -cK'_{i,j}\|\nabla\psi^t\| - K'_{i,j}(\epsilon\kappa_{i,j}^t\|\nabla\psi^t\|) + \nabla\psi_{i,j}^t \cdot \nabla K'_{i,j}, \quad (\text{A.13})$$

where Δt is the time step (e.g., Δt can be set to 0.05). The first term ($cK'_{i,j}\|\nabla\psi^t\|$) on the right hand side of the above equation is the velocity term (advection term) and, in the case of iris segmentation, acts as an inflation force. This term can lead to singularities and hence is discretized using upwind finite differences. The upwind scheme for approximating $\|\nabla\psi\|$ is given by:

$$\begin{aligned} \|\nabla\psi\| &= \sqrt{A}, \\ A &= \min(D_x^- \psi_{i,j}, 0)^2 + \max(D_x^+ \psi_{i,j}, 0)^2 + \\ &\quad \min(D_y^- \psi_{i,j}, 0)^2 + \min(D_y^+ \psi_{i,j}, 0)^2. \end{aligned}$$

where $D_x^- \psi$ is the first order backward difference of ψ in the x-direction; $D_x^+ \psi$ is the first order forward difference of ψ in the x-direction; $D_y^- \psi$ is the first order backward difference of ψ in the y-direction; and $D_y^+ \psi$ is the first order forward difference of ψ in the y-direction. The second term ($K'_{i,j}(\epsilon \kappa_{i,j}^t \|\nabla \psi^t\|)$) is a curvature based smoothing term and can be discretized using central differences. In our implementation, $c = 0.65$ and $\epsilon = 1$ for all iris images. The third geodesic term ($\nabla \psi_{i,j}^t \cdot \nabla K'_{i,j}$) is also discretized using the central differences.

After evolving the embedding function ψ according to Equation (A.13), the curve begins to grow until it satisfies the stopping criterion defined by the stopping function K' . But at times, the contour continues to evolve in a local region of the image where the stopping criterion is not strong. This leads to over-evolution of the contour. This can be avoided by minimizing the thin plate spline energy of the contour. By computing the difference in energy between two successive contours, the evolution scheme can be regulated. If the difference between the contours is less than a threshold (indicating that the contour evolution has stopped at most places), then the contour evolution process is terminated. The evolution of the curve and the corresponding embedding functions are illustrated in Figure A.6.

Since the radial fibers may be thick in certain portions of the iris, or the crypts present in the ciliary region may be unusually dark, this can lead to prominent edges in the stopping function. If the segmentation technique is based on parametric curves, then the evolution of the curve might terminate at these local minima. However, geodesic active contours are able to split at such local minima and merge again. Thus, they are able to effectively deal with the problems of local minima, thereby ensuring that the final contour corresponds to the true limbus boundary (Figure A.7).

A.2.2 Variational Level Sets

Another approach that can be used to precisely determine the iris boundaries is based on variational level sets [14]. This approach uses partial differential equations (PDE) to numerically solve the evolution of the curves that define the iris boundaries. The iris boundaries are first approximated using elliptical models, which are then refined using geometric active contours with variational formulation.

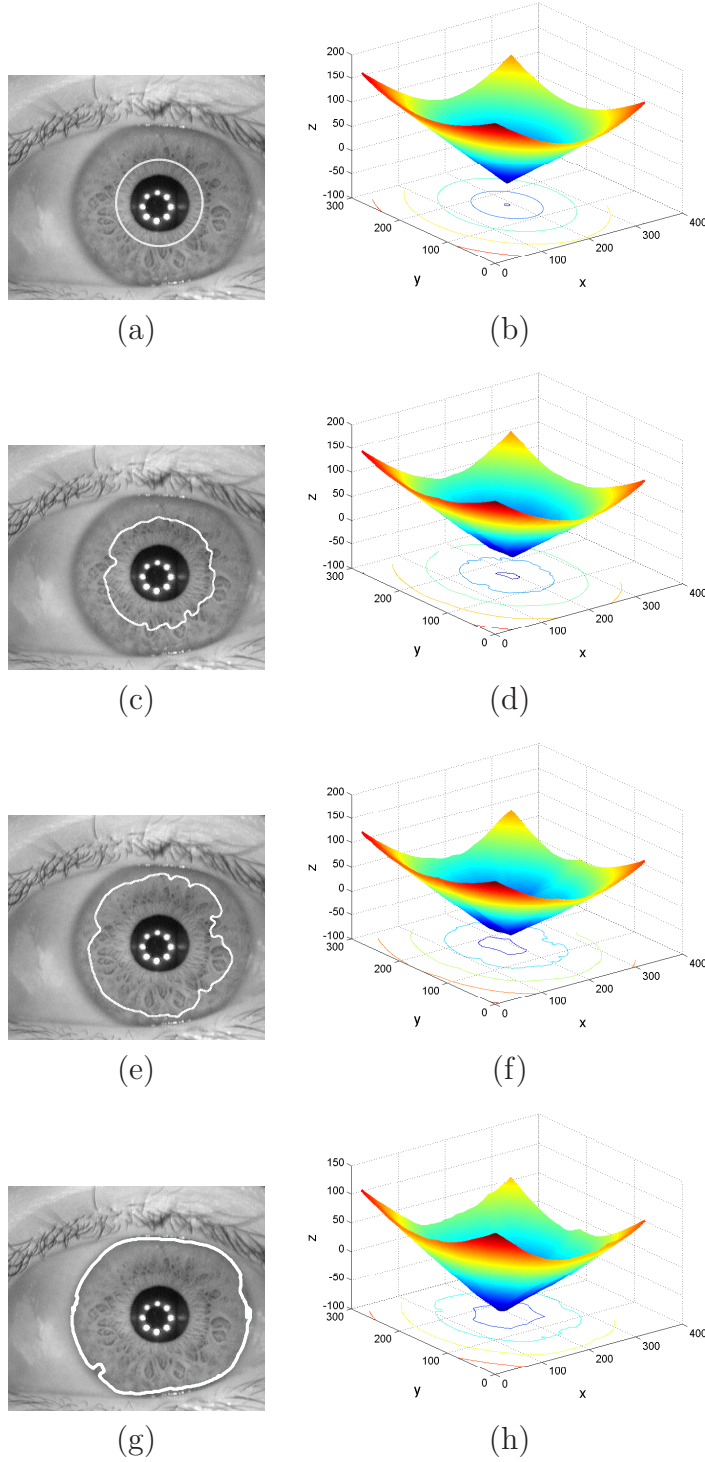


Figure A.6: Evolution of the geodesic active contour during iris segmentation. (a) Iris image with initial contour, (b) embedding function ψ (X and Y axes correspond to the spatial extent of the eye image and the Z axis represents different level sets), (c,d,e,f) contours after 600 and 1400 iterations, and their corresponding embedding functions, and (g,h) Final contour after 1800 iterations (contours shown in white).

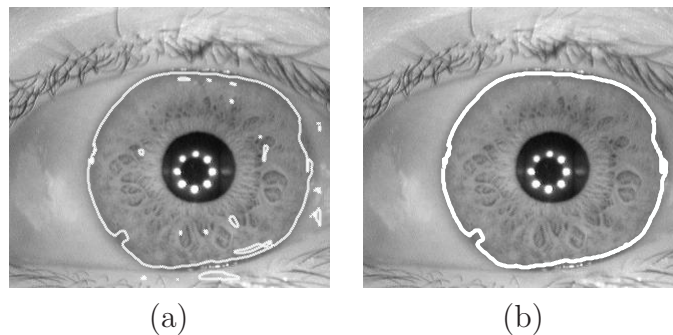


Figure A.7: The final contour obtained when segmenting the iris using the GAC scheme. (a) Example of a geodesic contour splitting at various local minima, (b) final contour (contours shown in white).

Given an iris image, first an elliptical model with parameters $(p_1, p_2, r_1, r_2, \varphi)$ is used to roughly determine the pupillary boundary. Here, (p_1, p_2) represents the center of the ellipse; r_1, r_2 , denote the semi-major axis and semi-minor axis, respectively; and φ denotes the orientation of the ellipse. By limiting the values of the semi-major and semi-minor axes to a specified range, the other parameters are iteratively varied with a small step size of three pixels to increase the size of the ellipse. At every iteration, a fixed number of points are randomly chosen on the circumference of the ellipse, and the total intensity difference between the chosen points and the center of the ellipse is computed. The boundary with the maximum intensity variation is chosen as the pupillary boundary. A rough contour of the limbus boundary is also determined in the same manner, with a different set of parameters for the semi-major and semi-minor axes. Figure A.8 shows an iris image with the rough contours of the pupillary and limbus boundaries obtained using the elliptical model.

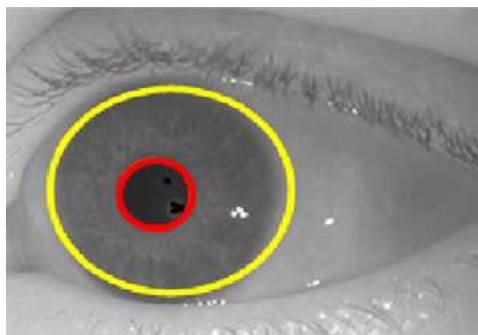


Figure A.8: Rough contours of the pupillary and limbus boundaries obtained using the elliptical model. Notice that the rough contours do not precisely match the true iris boundaries. Image source: Roy et. al. [14] ©Elsevier

Once the rough contours of the iris boundaries are obtained, the level set approach is applied to determine the precise contours. In the level set approach, the rough contour determined using the elliptical model is used as an active contour C , which can be represented as the zero level set $C(t) = \{(x, y) \mid \phi(t, x, y) = 0\}$ of a level set function $\phi(t, x, y)$. The objective of the level set approach is to determine a curve within the level set function, such that the total energy of the curve is minimum. The total energy of the curve is defined by the equation:

$$\varepsilon(\phi) = \mu\rho(\phi) + \varepsilon_{g,\lambda,\nu}(\phi), \quad (\text{A.14})$$

where the parameters $\varepsilon_{g,\lambda,\nu}(\phi)$ and $\rho(\phi)$ denote the external, and the internal energies of the curve, respectively, and $\mu > 0$. The external energy parameter depends on the image data, and drives the zero level set towards the boundary of the desired contour. On the other hand, the internal energy parameter helps in penalizing the deviation of the level set function, ϕ , from the signed distance function during the evolution of the curve.

The internal energy term in the above equation is further defined as:

$$\rho(\phi) = \int_{\Omega} \frac{1}{2} (|\nabla\phi| - 1)^2 dx dy, \quad (\text{A.15})$$

where Ω represents the image domain.

Similarly, the external energy term $\varepsilon_{g,\lambda,\nu}(\phi)$ can be further defined as:

$$\varepsilon_{g,\lambda,\nu}(\phi) = \lambda L_g(\phi) + \nu A_g(\phi), \quad (\text{A.16})$$

where $\lambda > 0$, and ν are constants. The term g denotes an edge detector function, and is defined as:

$$g = \frac{1}{1 + |\nabla G_{\sigma} * I|^2}, \quad (\text{A.17})$$

where G_{σ} denotes the Gaussian kernel with a standard deviation of σ , and I denotes the image. The term $L_g(\phi)$ is used to measure the length of the zero level set curve of ϕ , and is given by:

$$L_g(\phi) = \int_{\Omega} g \delta(\phi) |\nabla\phi| dx dy, \quad (\text{A.18})$$

where δ is the univariate Dirac function. The term $A_g(\phi)$ is used to speed up the curve evolution, and is defined as:

$$A_g(\phi) = \int_{\Omega} g H(-\phi) dx dy, \quad (\text{A.19})$$

where H is the Heaviside function.

The desired evolution equation of the level set function can be obtained by determining the value of $\frac{\partial \phi}{\partial t}$, using the following equation:

$$\frac{\partial \phi}{\partial t} = -\frac{\partial \varepsilon}{\partial \phi}, \quad (\text{A.20})$$

where $\frac{\partial \varepsilon}{\partial \phi}$ represents the Gateaux derivative of ε . The value of ϕ that minimizes the total energy function can be determined by satisfying the Euler-Lagrange equation, $\frac{\partial \varepsilon}{\partial \phi} = 0$. The Gateaux derivative of the functional ε can be written as follows:

$$\frac{\partial \varepsilon}{\partial \phi} = -\mu \left[\Delta \phi - \text{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) \right] - \lambda \delta(\phi) \text{div} \left(g \frac{\nabla \phi}{|\nabla \phi|} \right) - \nu g \delta(\phi). \quad (\text{A.21})$$

Thus, the desired equation of the level set function can be defined as:

$$\frac{\partial \phi}{\partial t} = \mu \left[\Delta \phi - \text{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) \right] + \lambda \delta(\phi) \text{div} \left(g \frac{\nabla \phi}{|\nabla \phi|} \right) + \nu g \delta(\phi). \quad (\text{A.22})$$

The last two terms on the right hand side of the above equation represent the gradient flows of the energy functional. These terms help in driving the zero level curve towards the boundaries of the required surface. The Dirac function $\delta(x)$ in the above equation is defined as:

$$\delta_\epsilon(x) = \begin{cases} 0, & |x| > \epsilon \\ \frac{1}{2\epsilon} \left[1 + \cos\left(\frac{\pi x}{\epsilon}\right) \right], & |x| \leq \epsilon. \end{cases} \quad (\text{A.23})$$

For the active contour ϕ , the curve evolution process for the pupillary boundary is carried out within a small range of ± 10 pixels from the rough contour. For the limbus boundary, this range is increased to ± 20 pixels. The curve evolution process is carried out from the outside of the approximated pupil boundary to avoid the effect of specular reflections. On the other hand, the process is carried out from the inside of the approximated limbus boundary to reduce the effect of eyelids and the eyelashes. Figure A.9 shows the final output obtained by the variational level set approach, in which both the iris boundaries are precisely detected. Figure A.10 shows the output of iris segmentation using the variational level set approach on some non-ideal iris images.

A.2.3 Fourier-based Approximation

Daugman [15] suggested the use of the Fourier series approximation in order to deduce the boundaries of the iris. The benefit of such an approximation is that the resulting output satisfies the following expectations:

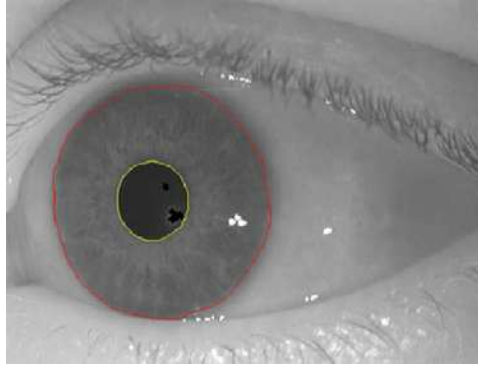


Figure A.9: Final output obtained using the variational level set approach, where both the iris boundaries are precisely determined. Image source: Roy et. al. [14] ©Elsevier

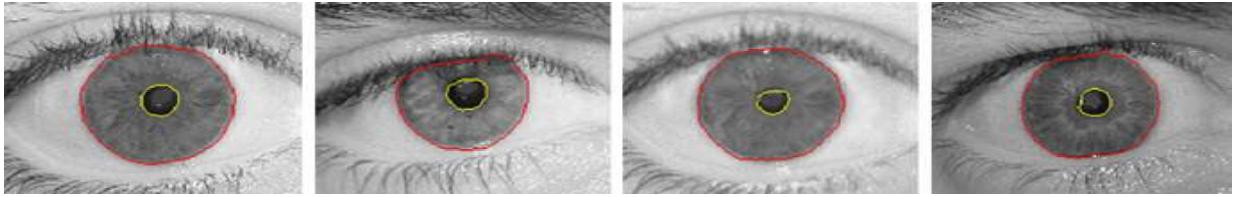


Figure A.10: Output of iris segmentation scheme using variational level set approach on some non-ideal iris images from the UBIRIS Version 1 dataset. Notice that the segmentation results are fairly precise. Image source: Roy et. al. [14] ©Elsevier

1. *Completeness*: An iris image may exhibit interruptions in its boundaries (e.g., interruptions caused by the specular reflections and eyelids to the pupillary and limbus boundaries, respectively). Therefore, the boundary detected by the segmentation algorithm must be robust to such interruptions.
2. *Closure*: Both the iris boundaries detected by the segmentation algorithm are expected to continue their trajectory across the interruptions on a principled basis, and form closed curves.

In this technique, given an image I , the coarse contour of the iris boundary is determined using active contours. Let the coarse iris contour be represented by N regularly spaced angular samples, given by $\{r_\theta\}$, $\theta = 0$ to $\theta = N - 1$. From this coarse contour, the corresponding iris boundary $\{R_\theta\}$, $\theta = 0$ to $\theta = N - 1$, that satisfies the above two conditions has to be determined. This can be achieved by the Fourier series approximation

of the coarse contour data, expressed as follows:

$$R_\theta = \frac{1}{N} \sum_{k=0}^{M-1} C_k \exp^{2\pi i k \theta / N}, \quad (\text{A.24})$$

where $\{C_k\}$ represents a set of M discrete Fourier coefficients, for $k = 0$ to $k = M - 1$, determined by the following equation:

$$C_k = \sum_{\theta=0}^{N-1} r_\theta \exp^{-2\pi i k \theta / N}. \quad (\text{A.25})$$

Generally, the zeroth-order Fourier coefficient (or the DC term C_0) describes the average curvature of the obtained boundary. Since this technique is used for determining the iris boundaries, the zeroth order coefficient determines the approximate radius of the output contour.

The value of M represents the number of active Fourier coefficients that are used in the approximation process. This value represents the number of degrees of freedom for the shape model of the boundary. In the most simplest scenario, i.e., $M = 1$, the model of the boundary will be circular. The value of M also acts as a tradeoff between the preciseness of the shape versus the strictness of the constraints (which corresponds to the complexity of the model). A strict set of constraints leads to a complex model, while a weak set of constraints leads to a simple model. Daugman suggests that $M = 17$ is a good choice for the pupillary boundary, and $M = 5$ for the limbus boundary. The two different choices for M are supported by the computer vision principle that strong data may be modeled with only weak constraints, while weak data should be modeled with strong constraints. The limbus boundary is considered as weak data because the occlusions caused by the eyelids and eyelashes are generally high. On the other hand, the pupillary boundary is considered as stronger data because the interruptions caused by the specular reflections are relatively minimal.

Figure A.11 shows the segmentation output obtained using the Fourier based approximation. The lower left corner of the image shows a *snake* that corresponds to the limbus boundary. The *snake* on the lower right corner of the image corresponds to the pupillary boundary. Both the snakes consist of two components: (a) a fuzzy ribbon-like data distribution that corresponds to the coarse contour, r_θ , and (b) a dotted curve that corresponds to the Fourier series approximation, R_θ . The characteristics of these snakes can be listed as follows:

1. The endpoints for both the snakes meet at the six o'clock position.
2. The thickness of each snake roughly corresponds to the sharpness of the corresponding edge.
3. The more un-interrupted an iris boundary is, the flatter and straighter the snake will be.

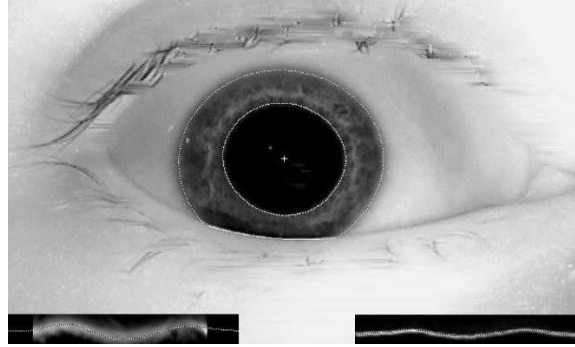


Figure A.11: Iris segmentation using Daugman's Fourier approximation approach. Image source: Daugman [15] ©IEEE. Image has been edited for clarity.

From the lower right corner of the figure, it can be noticed that the curve r_θ corresponding to the limbus boundary exhibits interruptions caused by the eyelid. However, the curve R_θ continues its trajectory even across the interruptions, proving the effectiveness of the approach.

A.3 Role of Image Acquisition on Iris Segmentation

Image acquisition plays an important role in the performance of an iris recognition system. Early iris recognition systems required significant cooperation from the subjects during image acquisition. This helped in acquiring good quality iris images, with minimal or no occlusions. As a trade-off, algorithms with low computational complexity [51, 62] were sufficient for the task of iris segmentation. With an increased demand for accurate iris recognition under practical scenarios (e.g., from a distance, under covert conditions, etc.), the requirements imposed during image acquisition are being relaxed. While this can impart flexibility to the image acquisition process, the quality of the acquired images can reduce drastically. In such cases, complex algorithms may be required to perform

segmentation while being robust to various non-ideal factors. This section discusses some iris segmentation algorithms, grouped according to the image acquisition conditions they are designed for.

Segmenting images acquired under highly constrained conditions

Images acquired under constrained conditions are often expected to be of high quality. This is because the user is typically still and cooperatively offers the iris images. Figure A.12 shows an image acquisition system which requires significant user cooperation. In such scenarios, the classical algorithms described in Section A.1 have been observed to provide good segmentation performance.



Figure A.12: Conventional iris image acquisition system requiring considerable user cooperation. Image source: <http://www.life.com/image/1668585>

Segmenting images acquired under less constrained conditions

The public usage iris recognition system proposed by Negin et. al. [16] may be considered to be one of the first systems that attempted to relax the image acquisition conditions. While other iris recognition systems prior to this work required the user to be in close proximity to the sensor, the proposed system allowed a stand-off distance of up to 3 feet. Furthermore, the proposed system allowed for an easy *public setup*, compared to its counterparts. An example of such a public setup could be the use of iris recognition technology to access an Automated Teller Machine (ATM). Even if the user at the ATM does not stand completely still, the system would be able to perform iris recognition during the transaction.

To use the system, the user was required to be reasonably cooperative and focus their gaze toward the system camera. The system would provide feedback to facilitate easy

image acquisition of the user's eye. A high-resolution video image of one eye of the user would be captured and used for recognition purposes. The image acquisition setup of the system is shown in Figure A.13.

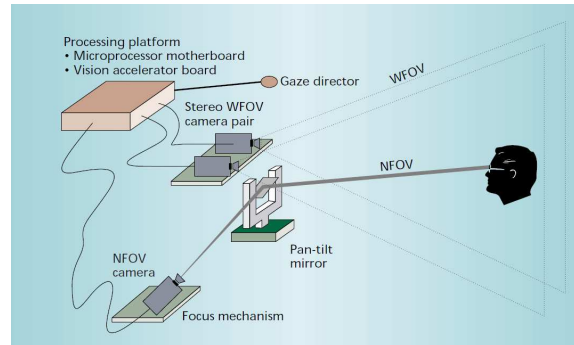


Figure A.13: Image acquisition setup for the public use iris recognition system. Image source: Negin et. al. [16] ©IEEE

The various steps involved in the working of this system, along with the functioning of the individual components, are provided below:

1. The user stands in front of the system, with a maximum allowable stand-off distance of 3 feet.
2. A wide field of view (WFOV) camera pair is used to capture an image of the user's torso, as shown in Figure A.14. The system then applies an image processing algorithm to locate the eyes of the user.



Figure A.14: Image of a user's torso acquired by the WFOV camera. Image source: Negin et. al. [16] ©IEEE

3. A narrow field of view (NFOV) camera pair is separately used to focus on the eye region. The system controls a gaze director to aid the user to look towards the

camera. A pan-tilt mirror is simultaneously used to direct the optical axis of the NFOV camera pair to ensure that the user's eye is focussed properly. As infrared illuminators are used during image acquisition, the system could operate even if the subject wore eyeglasses, contact lenses, or in a nighttime environment. Figure A.15 shows a sample image acquired at this stage.

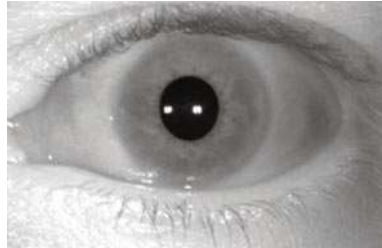


Figure A.15: Image of the user's eye, acquired by the NFOV camera. Image source: Negin et. al. [16] ©IEEE

4. A circular grid was used as a guide by the system to localize the iris region in the image acquired by the NFOV camera. The use of a circular grid simultaneously allowed for the exclusion of noisy regions such as the pupil, sclera, and the eyelids. The region lying within the grid was used for encoding and recognition. A sample image showing the circular grid for localizing the iris region is shown in Figure A.16.

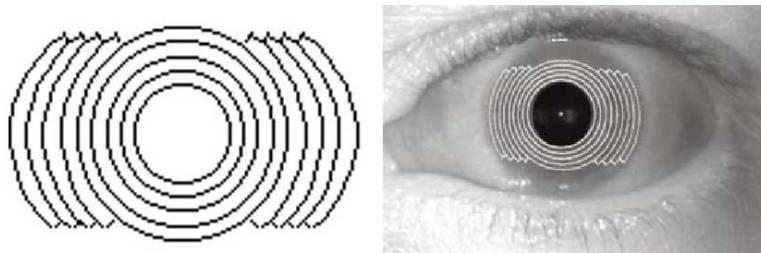


Figure A.16: Left: Circular grid used for iris localization. Right: Iris image with the circular grid overlaid. Image source: Negin et. al. [16] ©IEEE

While this system relaxed the acquisition conditions only moderately, it is considered to be significant in the field of iris biometrics for the following reasons:

1. This is one of the earliest works related to successful iris recognition from a distance.

2. This work highlights the dependency of iris segmentation on the eye detection scheme. When iris images are acquired from a distance, it has to be noted that eye detection has to be accurately performed to obtain good segmentation performance.

Segmenting images acquired using IOM (Iris On the Move) systems

The “Iris On the Move” (IOM) system developed by Matey et. al. [9] is considered to be another major development in the field of iris recognition from a distance. The IOM system significantly reduces the constraints on the position and the motion of a user during image acquisition. Such a flexibility is made possible by using an improved image acquisition system, that uses high-resolution cameras and video-synchronized illumination mounted on a minimally confining portal.

The setup of the IOM system consists of a walk through portal, similar to a metal detector. Near Infra-Red (NIR) illumination sources, and high-resolution image sensors are fixed to the portal. Images are acquired when a user walks through the portal at a normal walking speed ($< 1\text{ms}$). Stand-off distances up to 3 meters is possible, with a minimum requirement that the user be moderately cooperative. The system can acquire images even when a user wears eyeglasses or contact lenses, but cannot see through sunglasses. The camera used in the IOM system was Securimetrics PIER 2.3, which can acquire iris images with an approximate diameter of 200 pixels. As the heights of subjects can vary by a large factor, the system uses a set of cameras instead of one single camera. The setup of an IOM system is shown in Figure A.17.

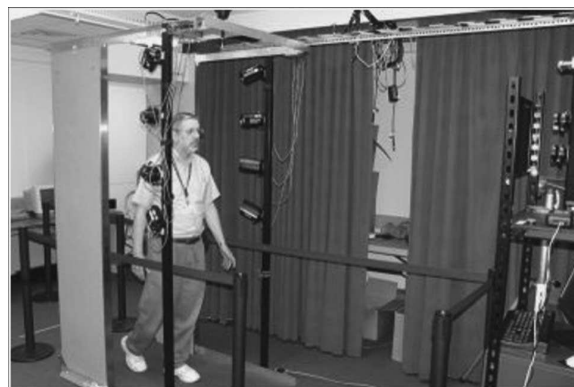


Figure A.17: Image acquisition setup for the iris on the move system. Image source: Matey et. al. [9] ©IEEE

The raw data acquired by the IOM system is typically a set of high-resolution facial images (2048×2048 pixels). This is possible due to the high-resolution and wide field-of-view of the cameras involved. The key requirement of the IOM system is that it should perform image acquisition, segmentation, and recognition in real time. The IOM system is considered to be an industrial application of the iris recognition technology that is expected to serve a large volume of people in short time. To reduce the computational time and processing speed of the system, the authors suggest a segmentation routine which is significantly different from the other segmentation schemes in the literature.

An iris image acquired by the IOM system typically exhibits a pattern of specular reflections on the iris. These specular reflections are caused by the Near Infra-Red (NIR) illumination system used during image acquisition. The pattern of the specular reflections are strongly dependent on the pattern in which the illuminators are arranged. Figure A.18 shows a sample NIR image exhibiting the specular reflections that correspond to the illumination pattern.

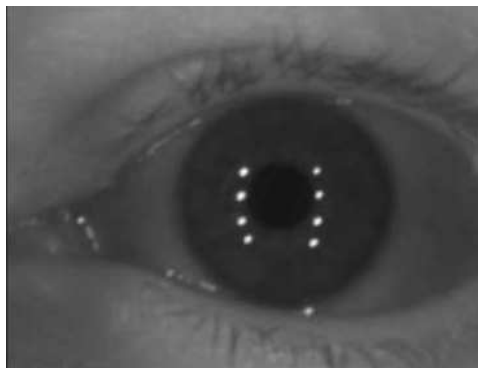


Figure A.18: Image acquired using the IOM system, exhibiting a specular reflection pattern that corresponds to the illuminator pattern. Image source: Matey et. al. [9] ©IEEE

Instead of ignoring specular reflections as noise, the segmentation scheme in the IOM system, in fact makes effective use of the pattern. A match filter is applied to the captured image producing the highest responses on the in-focus specularities. Once the specularities are highlighted, a thresholding scheme is used to binarize the image. This process yields an image showing only the locations of the specularities. The border, width, height and center of each specularity is determined in the binary image. As the illuminator pattern is pre-determined, specular patterns that are inconsistent with the illuminator pattern (or

those that are oddly shaped) are ignored. For the images that have strong correspondence between the specular reflection and the illuminator patterns, the iris regions around the specular reflections are extracted. Figure A.19 shows an iris image in which segmentation is performed using the specular reflection patterns.

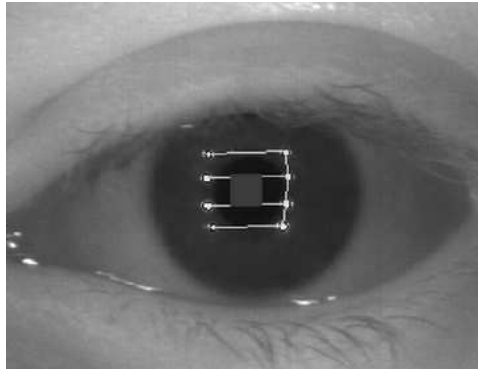


Figure A.19: Iris segmentation using the specular reflection pattern in the image. Once the specular reflection pattern is detected, a specified area around the pattern is used for unwrapping and feature extraction. Image source: Matey et. al. [9] ©IEEE

The contribution of the IOM system in the field of iris recognition is significant for the following reasons:

1. IOM system is the first image acquisition and recognition system that can work on iris images obtained from users walking at a normal pace.
2. The system allows for real-time recognition while relaxing the constraints imposed on the user, compared to its predecessors.

A.4 Segmenting Visible Wavelength Iris Images

Most iris recognition systems acquire input images using sensors that operate in the infra-red wavelength. This is due to the fact that the complex texture of the iris is more easily discernible when imaged under infra-red lighting. However, it may not be possible to acquire or operate with infra-red iris images at all times. Some of the reasons that support iris image acquisition in visible wavelength are provided below:

1. The current sensor and illuminator technology makes it difficult to acquire iris images from a distance using sensors that operate in the near infra-red wavelength.

2. Surveillance cameras usually work in visible wavelength. When using face and iris recognition systems on the data acquired from surveillance cameras, the iris data will be in the visible wavelength.

A majority of existing segmentation techniques use near infra-red images as their input. Such images typically exhibit higher contrast between the pupil and the iris regions, and induce the usual option of determining the pupillary border. In contrast, visible wavelength images usually exhibit less contrast between the pupil and the iris. This supports the inversion of the order in which the iris boundaries are segmented for visible images. An iris image acquired using a sensor that operates under visible wavelength is shown in Figure A.20.



Figure A.20: An iris image acquired in the visible wavelength. Notice that the intensity contrast between the pupil and the iris is low.

Proenca suggests an approach [78, 155] that performs automatic segmentation of the iris images acquired in the visible wavelength. Furthermore, the approach can perform iris segmentation on images that are acquired at a large imaging distance (4 to 8 meters), on the move, and under varying lighting. From the input image, the proposed technique first detects the sclera, as it is the most distinguishable region even under varying lighting conditions. Then, the fact that sclera mandatorily lies next to the iris is taken into account, and the iris regions are detected. A neural pattern recognition approach is later used to perform the segmentation.

The entire process can be broken down into three stages: detection of the regions that correspond to the iris; segmentation of the iris; and detection and elimination of the noisy regions. Given an image I , the sclera region is first detected. This is because in some images, iris detection is rather difficult. However, due to the naturally distinguishable appearance (color) of the sclera, it can be detected first by an approach which analyzes the color spaces of the image. Based on an empirical analysis, the author suggests three color

components: hue (h), blue (cb), and red chroma (cr), that can characterize information about the sclera. The contrast between the sclera and the remaining parts of the eye can be maximized using these three color components.

Given an iris image I , a 20-dimensional feature vector is constructed for each pixel, which can be denoted by the following expression:

$$\{x, y, h_{0,3,7}^{\mu,\sigma}(x, y), cb_{0,3,7}^{\mu,\sigma}(x, y), cr_{0,3,7}^{\mu,\sigma}(x, y)\} \quad (\text{A.26})$$

where x and y denote the position of the pixel, and $h()$, $cb()$, and $cr()$ denote the hue, blue, and red chroma components of the image at that pixel. The subscript denotes the radii of the circle that is centered at the pixel. The parameters μ , and σ , denote the mean, and the standard deviation, respectively, of the set of pixels which fall within those circular regions. For example, the term $cb_{0,3,7}^{\mu,\sigma}(x, y)$ means that six features were extracted from regions of the blue color component: three averages and three standard deviations, computed locally within regions of circles with radii 0, 3, and 7, centered at the considered pixel (x, y) . Once the feature vectors for all the image pixels are calculated for all the images, a neural network classifier is used to obtain the sclera map. This map indicates the location of sclera in the image.

To detect the iris, the information obtained using the sclera map is used. The author suggests that a pixel which lies within iris boundaries, when frontally imaged, will have similar number of sclera pixels on both sides. On the other hand, if the iris is off-angled, then the number of pixels of sclera on one side will be more than the number of pixels on the other side. To detect the iris pixels, another feature vector is generated. The data obtained from the sclera detection stage is used to obtain a new feature, referred to as “proportion of sclera” $p(x, y)$, for each image pixel. This feature helps in measuring the proportion of pixels that belong to the sclera in a direction d , with respect to the pixel at location (x, y) . The notation used for the directions are \uparrow for north, \downarrow for south, \leftarrow for east and \rightarrow for west. The feature vectors for each pixel are generated using the proportion of sclera information as follows:

$$p_{\leftarrow}(x, y) = \mu(sc((1, y - 1), (x, y))), \quad (\text{A.27})$$

$$p_{\rightarrow}(x, y) = \mu(sc((x, y - 1), (w, y))), \quad (\text{A.28})$$

$$p_{\uparrow}(x, y) = \mu(sc((x - 1, 1), (x, y))), \quad (\text{A.29})$$

$$p_{\downarrow}(x, y) = \mu(sc((x-1, y), (x, h))), \quad (\text{A.30})$$

where $sc((.,.), (.,.))$ denotes the regions of the image cropped from the detected sclera, delimited by the top-left and bottom-right corner coordinates. w and h denote the width and height, respectively. The value of $p()$ is set to 0 for all sclera pixels. The “proportion of sclera” values, pixel positions, local image saturation, and blue chrominance are then used to form a new feature vector represented as:

$$\{x, y, s_{0,3,7}^{\mu,\sigma}(x, y), cb_{0,3,7}^{\mu,\sigma}(x, y), p_{\leftarrow, \rightarrow, \uparrow, \downarrow}(x, y)\} \quad (\text{A.31})$$

where $s()$ and $cb()$ denote saturation and blue chrominance. Once again, the choice of color components is based on empirical evaluation.

Multilayered perceptron feed forward neural networks with one hidden layer are used by both classification stages. The neural network is trained using sample images and its output on a test image is considered to effectively decide the boundaries of the iris. Once the set of image pixels that correspond to a noise-free iris are identified, the goal is to determine the contours of the pupil and sclera of the iris. For this purpose, shape parameterization techniques are used.

A.5 Performance Evaluation of a Few Iris Segmentation Techniques

To provide an understanding of the iris segmentation performance, 3 techniques were chosen. The three algorithms are (a) integro-differential operator, (b) Hough transformation, and (c) Geodesic Active Contours. The segmentation performance is evaluated on the ICE database [104], which contains a total of 2953 iris images corresponding to both left and right eyes. The images are in the TIFF format, with a resolution of 640×480 pixels. The quality of the images is reasonably good, as they were collected from cooperative users. However, some images exhibit eyelid or eyelash occlusions. The performance of an iris segmentation technique was measured by computing the segmentation accuracy as follows:

$$\text{Segmentation accuracy} = \frac{\text{Number of correctly segmented images}}{\text{Number of input images}} \times 100 \quad (\text{A.32})$$

Table A.5 lists the segmentation accuracies of the three techniques. From the results, it can be observed that Geodesic Active Contours provide better segmentation performance

Segmentation technique	Number of input images	Number of correctly segmented images	Segmentation accuracy
Integro-differential operator	2953	2597	87.9%
Hough transform	2953	2632	89.1%
Geodesic Active Contours (GAC)	2953	2699	91.3%

Table A.1: **Segmentation accuracies of the three techniques evaluated in this work.**

when compared to the classical approaches. However, it has to be noted that the computational expense of the classical approaches is much lower than that of the Geodesic Active Contours. Thus, it is necessary to understand the computational demands and operational requirements of an application prior to engaging a specific iris segmentation algorithm. It may be possible to dynamically determine *which* iris segmentation algorithm is appropriate to be used based on characteristics of the image to be segmented. Further, the outputs of multiple segmentation algorithms may be combined to generate a single hypothesis for the boundaries of the iris.

A.6 Approaches to Refine Iris Segmentation

One of the major concerns in iris segmentation is the over- or under-segmentation of the iris boundaries. Over segmentation refers to the situation where the radius of the detected iris boundary is larger than that of the actual boundary. On the other hand, under segmentation refers to the situation where the radius of the detected iris boundary is smaller than the actual boundary. Figure A.21 shows a sample image for each case.

In both cases, the offset between the actual iris boundary and the detected iris boundary is not large. However, such minor offsets can significantly lower iris recognition performance. This is due to the inclusion of noise or regions that do not contain discriminatory texture information (e.g., sclera, eyelashes, etc.). In this section, some approaches are discussed that attempt to refine segmentation by operating on finer details in the vicinity of the iris.



Figure A.21: Left: Over segmented limbus boundary. Right: Under segmented limbus boundary.

Eyelash Removal in Iris Images

One of the many factors that affect iris recognition performance is the occlusion caused by eyelashes. While some approaches ignore the iris regions occluded by eyelashes [64], others detect the eyelashes and mask them while encoding [87]. On the contrary, the technique proposed by Zhang et. al. [17] attempts to restore the iris regions that are occluded by the eyelashes. A non-linear conditional directional filtering approach is used to perform the restoration. The proposed technique is an iterative approach involving the following steps: (a) detecting the pixels that correspond to eyelash occlusion, (b) detecting the direction of the eyelash that causes the occlusion, (c) local filtering of occlusion region in the direction that is perpendicular to the eyelash, (d) restoring the pixel intensity by using a 1D median filter.

The proposed technique is more suited for unwrapped or normalized iris images, rather than the original iris images. The unwrapped iris images in this work are of size 512×80 pixels, and the top 48 rows of pixels nearest to the pupil are used by the eyelash removal technique. Figure A.22 shows eyelash occlusion in an unwrapped iris image.

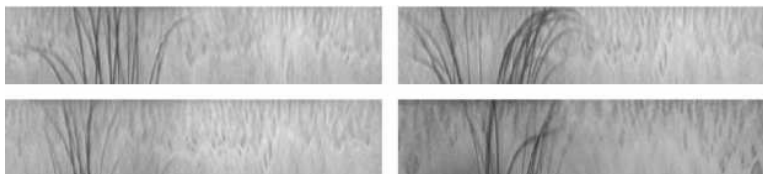


Figure A.22: Unwrapped iris images containing eyelash occlusion. Image source: Zhang et. al. [17] ©IEEE

An eyelash is expected to cause a discontinuity along its edges. Therefore, given an

unwrapped iris image I , an eyelash is detected, and its direction is estimated by using an edge filter. For this purpose, a 3×3 Sobel edge filter is applied on the image. A Sobel edge filter is shown in Figure A.23.

-1	-2	-1	z_1	z_2	z_3	-1	0	1
0	0	0	z_4	z_5	z_6	-2	0	2
1	2	1	z_7	z_8	z_9	-1	0	1

Figure A.23: Left: x derivative for a Sobel Edge filter. Center: image region under consideration. Right: y derivative for a Sobel edge filter.

For each pixel, the gradients in the x and y directions $[G_x, G_y]$ can be determined by the following equations:

$$G_x = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3) \quad (\text{A.33})$$

$$G_y = (z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7) \quad (\text{A.34})$$

Similarly, the magnitude of the gradient at the center of the mask, $Grad$, is computed as:

$$Grad = \sqrt{(G_x^2 + G_y^2)}. \quad (\text{A.35})$$

The local gradient direction that is perpendicular to the edge can then be determined by:

$$\theta = \arctan(G_y/G_x). \quad (\text{A.36})$$

To determine if a pixel is occluded or not, a window of size $m \times n$ is centered at a pixel, and the gradient direction variance for the r pixels that lie within the window and have a $Grad$ value above a specific threshold is computed as follows:

$$Var - Grad = \frac{1}{r-1} \sum_{i=1}^r (\theta_i - \bar{\theta})^2. \quad (\text{A.37})$$

A strong edge is indicated if the gradient direction has a small variance, and the pixel is classified as being affected by an eyelash. To restore such a pixel, a 1D median filter of length L is applied along the direction θ . This process outputs an estimate of the value of the image with the eyelash removed.

To avoid incorrectly filtering non-eyelash pixels, pixel alteration is carried out only if the change in the pixel exceeds a certain threshold, related to the total variance of the image. For this purpose, a parameter *Recover* is computed as follows:

$$Recover = Diff - k * Var(Image) \quad (A.38)$$

where *Diff* represents the difference in the intensity between the filtered and the unfiltered pixel, *Var(Image)* represents the intensity variance of the whole unfiltered image, *k* denotes parameter used to tune the threshold. The pixel is replaced by the filtered value, only if *Recover* is positive. By using this approach, the visual appearance is not significantly changed, but the recognition performance is improved. Figure A.24 shows an unwrapped iris image before and after applying the approach.



Figure A.24: Top: Image affected by eyelash occlusion. Bottom: Output obtained by applying the eyelash removal technique. Image source: Zhang et. al. [17] ©IEEE

Improving Daugman's Classical Segmentation Algorithm

Libor Masek's MATLAB package [77] is one of the most widely used open source implementation for iris segmentation. For good quality iris images acquired under regular imaging conditions, Libor Masek's implementation results in good segmentation performance. However, Liu et. al. [18] showed that the segmentation performance can be further improved by incorporating minor modifications to Libor Masek's implementation. The two most significant modifications that were used to improve the segmentation performance of Libor Masek's implementation are as follows:

1. Reversal of the detection order of the iris boundaries: In Masek's implementation, the limbus boundary is detected first, followed by the pupillary boundary. However, by reversing the detection order, slightly better segmentation performance can be observed. This is based on the fact that the pupillary boundary exhibits strong intensity variation at its boundary, when compared to that of the limbus boundary. By reversing the order, the dependency problem can be minimized.

2. Eyelid detection: Libor Masek’s implementation models the eyelids as two horizontal lines. As a result of such an approximation, some of the iris texture can be occluded. The authors suggest splitting the top and bottom eyelid regions into two different portions each as shown in Figure A.25, and then performing eyelid detection. As a result of this, the eyelids will no longer be approximated as straight lines but will appear curvy. This modification avoids unnecessary occlusion of the iris during eyelid estimation. Figure A.26 shows the difference observed in eyelid detection using the proposed modification.

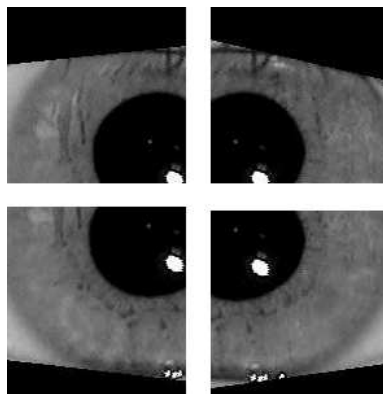


Figure A.25: Splitting of the eyelid into four portions to allow better detection of eyelid. Image source: Liu et. al. [18] ©IEEE

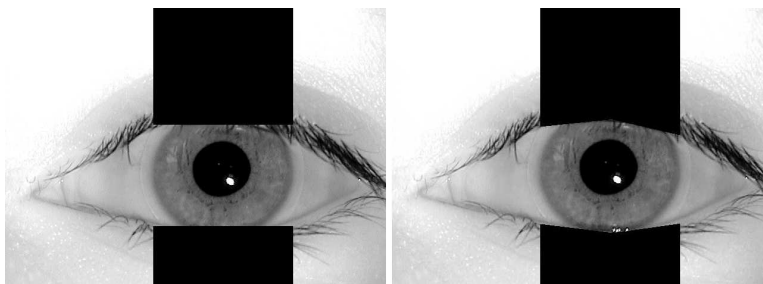


Figure A.26: Left: Eyelid detection using Libor Masek’s approach. Right: Improved eyelid detection using the proposed modifications. Image source: Liu et. al. [18] ©IEEE

A.7 Predicting Errors in Iris Segmentation

The performance of an iris recognition system is highly dependent on the output of iris segmentation. If an incorrectly segmented region is used for recognition, the lack of rich distinctive texture can reduce recognition performance. Thus, designing an algorithm

that can examine the input image and predict in advance if segmentation is likely to fail or not would be beneficial. Such a scheme can be used to provide feedback during image acquisition, requiring the user to provide a better, more useful image. When acquiring a new image is not a possibility, the image can at least be flagged to avoid being used for recognition purposes. Manual segmentation could then be used to further process such an image. Another possibility is to design an algorithm that can evaluate the output of the segmentation routine and determine if the segmentation has failed or not.

Some algorithms compute the quality of the iris region using local image analysis [58]. However, such algorithms require at least a coarsely segmented iris. If the segmentation is improper, the quality estimate will be incorrect, thus defeating the motive for such an exercise.

Kalka et. al. [75] propose an algorithm which evaluates the output of the pupil and iris segmentation routines. The algorithm is based on combining probabilistic intensity features with geometric features to generate scores that indicate the success of segmentation. A decision tree-based machine learning approach is used to render a binary decision: success or failure. The proposed method, therefore, predicts whether the output of segmentation is good (both the pupil and iris boundaries were correctly estimated) or bad (at least one of the boundaries was incorrectly estimated).

The following measures are taken into account:

1. Pupil segmentation measure: One of the major concerns here is related to the over- or under-segmentation of the pupil. To take into account such problems, a probabilistic model is used to fit the segmentation output for the pupil. The output of the model is then used to generate an over-segmentation or under-segmentation score.
2. Geometric iris measure: This is based on the fact that the limbus and pupillary boundaries are actually concentric (circles or ellipses) when they are frontally imaged. Therefore, a measure based on eccentricity and concentricity of the boundaries is taken into account.

Given an iris image segmentation output, the goal is to assign a score to that output, which indicates the “goodness” of segmentation based on the above factors. First, the pupil boundary is checked to establish whether it is over- or under-segmented. For this, consider

an image I , and let its pupillary contour be represented as \bar{x} . The task is to decide whether the pixels lying within the boundary \bar{x} actually belong to the pupil or not. To this end, a probabilistic model is associated with \bar{x} , which formulates a likelihood ratio test, $\Lambda(\bar{x})$, as follows:

$$\Lambda(\bar{x}) = \frac{P(\bar{x} | H_1)}{P(\bar{x} | H_0)} \geq \eta \quad (\text{A.39})$$

where $H_1 : \bar{x}$ corresponds to a pupil pixel and $H_0 : \bar{x}$ corresponds to a non-pupil pixel. Based on empirical evaluation, the authors use a Gamma distribution $\Gamma(k, \Theta)$, for $P(\bar{x} | H_1)$, and a Gaussian distribution $N(\mu, \sigma^2)$ for $P(\bar{x} | H_0)$, respectively. The value of the shape parameter k is fixed to 1. To obtain the scale parameter value Θ , the spatial histogram of the image intensities in the region of interest is computed. From the histogram, the scale parameter is estimated as:

$$\hat{\Theta} = \sum_{i=0}^{Bin(P_t)} x_i w_i \quad (\text{A.40})$$

where P_t denotes the threshold used to constrain the size of the pupil region, x_i is a gray level bin from the histogram of the region of interest, and w_i is the weight associated with bin x_i . In other words, Θ is obtained by summing the product of the gray level bins and the associated weights until the bin corresponding to P_t is reached. The parameters of the Gaussian are estimated using the following equations:

$$\hat{\mu} = \sum_{i=Bin(P_t)+1}^{Bin(I_t)} x_i w_i \quad (\text{A.41})$$

$$\hat{\sigma}^2 = \sum_{i=Bin(P_t)+1}^{Bin(I_t)} w_i (x_i - \hat{\mu})^2 \quad (\text{A.42})$$

where I_t is a threshold used to constrain the size of the iris region, x_i is the gray level bin in the histogram corresponding to the region of interest, and w_i is the weight associated with bin x_i . Thresholds P_t and I_t are determined experimentally. Every pixel within the pupil boundary is assigned a 0 or 1, based on the likelihood ratio test. Once values are assigned to all pixels in the region of interest, the over-segmentation measure P_{over} is computed as the ratio of the number of 0's to that of 1's.

To measure the under-segmentation of the pupil boundary, an iterative approach is employed where the estimated pupil radius (or major and minor axes in case of an ellipse) is increased and it is determined whether the pixels inside the expanded region are pupil

pixels by the same approach as above. This process is continued until the pupil radius reaches the size of the iris radius, or the ratio of pupil to non-pupil pixels is less than 20%. The use of a threshold prevents the influence of heterogenous factors such as dark eyelashes/eyelids and reduces unnecessary computations. The final under segmentation score is defined as follows:

$$P_{under} = \frac{P_{over}}{P_{over} + P_{est_under}} \quad (\text{A.43})$$

P_{est_under} is the total number of estimated pupil pixels over all iterations.

To obtain yet another score value for the accuracy of segmentation, the concentricity and eccentricity of the iris and pupil boundaries are used. These values are measured using the following equations:

$$I_C = \sqrt{(p_x - i_x)^2 + (p_y - i_y)^2} \quad (\text{A.44})$$

$$I_E = \sqrt{(p_x - i_x)^2 + (p_y - i_y)^2} + \arccos\left(\frac{b_i}{a_i}\right) * 100 + \arccos\left(\frac{b_p}{a_p}\right) * 100 \quad (\text{A.45})$$

where (p_x, p_y) are the pupil center coordinates, (i_x, i_y) are the iris center coordinates, b_i and a_i are the semi-minor and semi-major axes for the iris ellipse, respectively, and b_p and a_p are the semi-minor and semi major axes for the pupil ellipse, respectively.

Once both the pupil segmentation score and the iris segmentation scores are available, a Naive-Bayes Tree classifier is used to generate a final score. This final score is binarized to indicate the performance of iris segmentation.

Appendix B

Database Release Agreement

STUDENT DATA USAGE REQUIREMENTS AND RESTRICTIONS

West Virginia University

Biometric Collection of People Dataset Release Agreement

Modality: ☒ Face ☐ Fingerprint ☒ Iris ☐ Latent Fingerprint

Release #: BioCoP 1

Enclosed is a government furnished information (GFI) agreement. All GFI is governed by Title 28 of the Code of Federal Regulations (CFR) § 22. Please read the governing section of code as well as the enclosed agreement and return a signed copy indicating your intent to be bound by the rules put forth by the federal government and this agreement. All partners in research intending to use the GFI must use it in accordance with 28 CFR § 22.

In addition to those duties and requirements outlined in the attached GFI Agreement and 28 CFR § 22, student use of the biometric dataset is further restricted by the following:

- The entire or part of the biometric dataset will not be further distributed, copied or disseminated on any electronic or storage device to be taken outside of the secure environment in which the dataset is stored. Data will be housed and used only on designated in-house devices and will not be removed from the secure premises in any form.

I have read and accept these provisions as conditions for the receipt and use of the GFI. I agree to comply with these provisions throughout the lifetime of the Cooperative Agreement. I have read this Agreement, restrictions outlined in this document, and the applicable Code of Federal Regulations carefully and questions, if any, have been answered.

RAGHAVENDER R. JILLELA

Name (Printed)

 April 18, 2012
Signature and DateWEST VIRGINIA UNIVERSITY

Organization

ARUN ROSS

Name of Supervising Personnel

 4/18/2012
Signature and Date

Bibliography

- [1] O. V. Komogortsev, A. Karpov, L. R. Price, and C. Aragon, “Biometric Authentication via Oculomotor Plant Characteristics,” in *5th IAPR International Conference on Biometrics*, April 2012, pp. 413–420.
- [2] Wikimedia Commons, “Schematic diagram of the human eye,” <http://upload.wikimedia.org/wikipedia/commons/1/1e/Schematic/diagram/of/the/human/eye/en.svg>.
- [3] S. Manvelyan, “Close-up images of the human iris,” <http://www.surenmanvelyan.com>.
- [4] A. Ross, R. Jillela, J. Smereka, V. Boddeti, B. VijayaKumar, R. Barnard, X. Hu, P. Pauca, and R. Plemmons, “Matching Highly Non-ideal Ocular Images: An Information Fusion Approach,” *Proceedings of the Fifth IAPR International Conference on Biometrics*, March 2012.
- [5] “Congenital Iris Abnormality,” <http://sciencero11.files.wordpress.com/2007/06/congenitalirisabnorm.jpg>.
- [6] S. Baker, K. Bowyer, and P. Flynn, “Empirical Evidence for Correct Iris Match Score Degradation with Increased Time Lapse between Gallery and Probe Images,” in *International Conference on Biometrics*, 2009, pp. 1170–1179.
- [7] NIST, “Face and Ocular Challenge Series Database,” <http://www.nist.gov/itl/iad/ig/focs.cfm/>, 2011.
- [8] NIST, “NIST Test Proves ‘The Eyes Have It’ for ID Verification,” <http://www.nist.gov/itl/csd/id/110309.cfm>.
- [9] J. R. Matey, O. Naroditsky, K. Hanna, R. Kolczynski, D. J. LoIacono, S. Mangru, M. Tinker, T. M. Zappia, and W. Y. Zhao, “Iris on the Move: Acquisition of Images for Iris Recognition in Less Constrained Environments,” *Proceedings of the IEEE*, vol. 94, no. 11, pp. 1936–1947, 2006.
- [10] Sarnoff, “Iris On Move,” <http://www.sarnoff.com/demos/iris-on-the-move/>, 2012.
- [11] Subject ID: 02463d3176, “BEST Development Challenge Program (BDGP) Database,” 2012.
- [12] “Fake iris,” <http://pubs.acs.org/cen/img/83/i41/8341sci2eyelive1.jpg/>.

- [13] NIST, “Face Recognition Grand Challenge Database,” <http://www.frvt.org/FRGC/>.
- [14] K. Roy, P. Bhattacharya, and C. Suen, “Iris Segmentation using Variational Level Set Method,” *Optics and Lasers in Engineering*, vol. 49, no. 4, pp. 578–588, 2011.
- [15] J. Daugman, “New Methods in Iris Recognition,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 37, no. 5, pp. 1167–1175, 2007.
- [16] M. Negin, T. Chmielewski, M. Salganicoff, U. von Seelen, P. Venetainer, and G. Zhang, “An Iris Biometric System for Public and Personal Use,” *Computer*, vol. 33, no. 2, pp. 70–75, February 2000.
- [17] D. Zhang, D. Monroe, and S. Rakshit, “Eyelash Removal Method for Human Iris Recognition,” in *IEEE International Conference on Image Processing*. IEEE, 2006, pp. 285–288.
- [18] X. Liu, K. Bowyer, and P. Flynn, “Experiments with an Improved Iris Segmentation Algorithm,” in *IEEE Workshop on Automatic Identification Advanced Technologies*. IEEE, 2005, pp. 118–123.
- [19] Federal Bureau of Investigation (FBI) and West Virginia University (WVU), “Biometric Collection of People (BioCoP) database,” Proprietary multi-modal biometric database, 2012.
- [20] D. Usher, Y. Tosa, and M. Friedman, “Ocular Biometrics: Simultaneous Capture and Analysis of the Retina and Iris,” in *Advances in Biometrics*, N. K. Ratha and V. Govindaraju, Eds., pp. 133–155. Springer London, 2008.
- [21] L. Flom and A. Safir, “U. S. Patent No. 4 641 349,” U. S. Government Printing Office, Washington, D. C., 1987.
- [22] R. Derakhshani and A. Ross, “A Texture-Based Neural Network Classifier for Biometric Identification using Ocular Surface Vasculature,” in *International Joint Conference on Neural Networks*, August 2007, pp. 2982–2987.
- [23] D. Woodard, S. Pundlik, P. Miller, R. Jillela, and A. Ross, “On the Fusion of Periocular and Iris Biometrics in Non-Ideal Imagery,” in *20th International Conference on Pattern Recognition*. IEEE, 2010, pp. 201–204.
- [24] U. Park, R. Jillela, A. Ross, and A. K. Jain, “Periocular Biometrics in the Visible Spectrum,” *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 1, pp. 96–106, March 2011.
- [25] P. Kasrowski and J. Ober, “Eye Movements in Biometrics,” in *Biometric Authentication*, D. Maltoni and A. K. Jain, Eds., vol. 3087 of *Lecture Notes in Computer Science*, pp. 248–258. Springer Berlin / Heidelberg, 2004.
- [26] C. Holland and O. Komogortsev, “Biometric Verification via Complex Eye Movements: The Effects of Environment and Stimulus,” in *IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*. IEEE, 2012, pp. 39–46.

- [27] O. H. Perry Pepper, "Medical Etymology: The History and Derivations of Medical Terms for Students of Medicine, Dentistry, and Nursing," Last accessed on October 1, 2012, from <http://www.emory.edu/ANATOMY/AnatomyManual/Etymology.html>.
- [28] D. Harper, "Online Etymology Dictionary," Last accessed on October 1, 2012, from <http://www.etymonline.com/>.
- [29] F. H. Adler, "Physiology of the Eye: Clinical Application," *London: The C. V. Mosby Company*, 1965.
- [30] H. Davson, "Physiology of the Eye," *Elsevier Science Limited, 5th edition*, 1990.
- [31] P. C. Kronfeld, "Gross Anatomy and Embryology of the Eye," in *The Eye*, H. Davson, Ed. London: Academic Press, 1962.
- [32] M. Sam Sunder and A. Ross, "Iris Image Retrieval Based on Macro-features," in *Proceedings of the 20th International Conference on Pattern Recognition*. IEEE, 2010, pp. 1318–1321.
- [33] J. G. Daugman, "High Confidence Visual Recognition of Persons by a Test of Statistical Independence," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, pp. 1148–1161, 1993.
- [34] A. Bertillon, "La couleur de l'iris," *Revue Scientifique*, vol. 36, no. 3, pp. 65–73, 1885.
- [35] A. Bertillon, "Alphonse Bertillon's Instructions for taking descriptions for the Identification of Criminals, and others, by the means of Anthropometric Indications," in *Identification anthropometrique, instructions signaltiques, Translated from the French by Muller, G.* American Bertillon Prison Bureau, Chicago, IL, 1889.
- [36] G. Davenport and C. Davenport, "Heredity of Eye-Color in Man," *Science*, vol. 26, pp. 589–592, 1907.
- [37] I. Mann, "The Development of the Human Eye," *Cambridge University Press, for the British Journal of Ophthalmology*, p. 306, 1928.
- [38] J. Rohen and H. Unger, "Studies on the Morphology and Pathology of the Trabecular Meshwork in the Human Eye," *American Journal of Ophthalmology*, vol. 46, no. 6, pp. 802–813, 1958.
- [39] Wasserman. H. P., "Ethnic Pigmentation: Historical, Physiological, and Chemical Aspects," *American Elsevier, New York*, 1974.
- [40] R. S. Worrall, "Iridology: Diagnosis or Delusion," *The Skeptical Inquirer*, pp. 23–35, 1983.
- [41] L. Berggren, "Iridology: A Critical Review," *Acta Ophthalmologica*, vol. 63, no. 1, pp. 1–8, 1985.
- [42] F. W. Newell, "Ophthalmology Principles and Practice, 7th edition," *Mosby, St. Louis, MO*, 1991.

- [43] R. G. Johnson, "Can Iris Patterns be Used to Identify People?, Technical Report LA-12331-PR," Tech. Rep., Los Alamos National Laboratory, CA, Chemical and Laser Sciences Division, 1991.
- [44] R. Johnston and W. Grace, "Bartus Iris System, Technical Report LA-UR-96-1969," Tech. Rep., Los Alamos National Laboratory, 1996.
- [45] F. Bouchier, J. Ahrens, and G. Wells, "Laboratory Evaluation of the IriScan Prototype Biometric Identifier, Technical Report SAND96-1033, UC-900," Tech. Rep., Sandia National Laboratories, Albuquerque, NM., 1996.
- [46] J. Zuo, N. Kalka, and N. Schmid, "A Robust Iris Segmentation Procedure for Unconstrained Subject Presentation," in *2006 Biometrics Symposium: Special Session on Research at the Biometric Consortium Conference*, August 2006, pp. 1–6.
- [47] K. W. Bowyer, K. Hollingsworth, and P. J. Flynn, "Image Understanding for Iris Biometrics: A Survey," *Computer Vision and Image Understanding*, vol. 110, no. 2, pp. 281–307, 2008.
- [48] J. Daugman, "Probing the Uniqueness and Randomness of IrisCodes: Results from 200 Billion Iris Pair Comparisons," *Proceedings of the IEEE*, vol. 94, no. 11, pp. 1927–1935, 2006.
- [49] C. Belcher and Y. Du, "Region-based SIFT Approach to Iris Recognition," *Optics and Lasers in Engineering*, vol. 47, no. 1, pp. 139–147, 2009.
- [50] J. van der Gracht, V.P. Pauca, H. Setty, R. Narayanswamy, R. Plemmons, S. Prasad, and T. Torgersen, "Iris Recognition with Enhanced Depth-of-Field Image Acquisition," *Proceedings of SPIE*, vol. 5438, pp. 120–129, 2004.
- [51] R. Wildes, "Iris Recognition: An Emerging Biometric Technology," *Proceedings of the IEEE*, vol. 85, no. 9, pp. 1348–1363, September 1997.
- [52] K. Smith, V. Pauca, A. Ross, T. Torgersen, and M. King, "Extended Evaluation of Simulated Wavefront Coding Technology in Iris Recognition," in *IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, 2007, pp. 1–7.
- [53] R. Plemmons, M. Horvath, E. Leonhardt, V. Pauca, S. Prasad, S. Robinson, H. Setty, T. Torgersen, J. van der Gracht, E. Dowski, R. Narayanswamy, and P. Silveira, "Computational Imaging Systems for Iris Recognition," in *Proceedings of the SPIE*, 2004, vol. 5559, pp. 346–357.
- [54] S. Tankasala, R. Pasula, V. Gottemukkula, S. Saripalle, V. Nalamati, A. Ross, and R. Derakhshani, "A Video based Hyper-Focal Imaging Method for Iris Recognition in the Visible Spectrum," *IEEE International Conference on Technologies for Homeland Security*, November 2012.
- [55] C. Boyce, A. Ross, M. Monaco, L. Hornak, and X. Li, "Multispectral Iris Analysis: A Preliminary Study," in *Computer Vision and Pattern Recognition Workshop*, June 2006.

- [56] A. Ross, R. Pasula, and L. Hornak, "Exploring Multispectral Iris Recognition Beyond 900nm," in *IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems (BTAS)*, September 2009, pp. 1–8.
- [57] R. Ives, H. Ngo, S. Winchell, and J. Matey, "Preliminary Evaluation of Multispectral Iris Imagery," in *IET Conference on Image Processing*, July 2012, pp. 1–5.
- [58] Y. Chen, S. Dass, and A. Jain, "Localized Iris Image Quality using 2D Wavelets," *Advances in Biometrics*, pp. 373–381, 2005.
- [59] N. Kalka, J. Zuo, V. Dorairaj, N. Schmid, and B. Cukic, "Image Quality Assessment for Iris Biometric," *Proceedings of 2006 SPIE Conf. on Biometric Technology for Human Identification III*, vol. 6202, pp. 61020D–1–62020D–11, April 2010.
- [60] N. Kalka, J. Zuo, N. Schmid, and B. Cukic, "Estimating and Fusing Quality Factors for Iris Biometric Images," *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, vol. 40, no. 3, pp. 509–524, May 2010.
- [61] X. Li, Z. Sun, and T. Tan, "Comprehensive Assessment of Iris Image Quality," in *18th IEEE International Conference on Image Processing*, September 2011, pp. 3117–3120.
- [62] J. Daugman, "How Iris Recognition Works," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 21–30, January 2004.
- [63] L. Ma, Y. Wang, and T. Tan, "Iris Recognition using Circular Symmetric Filters," in *Proceedings of the 16th IEEE International Conference on Pattern Recognition*, 2002, vol. 2, pp. 414–417.
- [64] L. Ma, T. Tan, Y. Wang, and D. Zhang, "Efficient Iris Recognition by Characterizing Key Local Variations," *IEEE Transactions on Image Processing*, vol. 13, no. 6, pp. 739–750, 2004.
- [65] J. Thornton, M. Savvides, and B. Vijayakumar, "Robust Iris Recognition using Advanced Correlation Techniques," *Image Analysis and Recognition*, pp. 1098–1105, 2005.
- [66] J. Huang, Y. Wang, T. Tan, and J. Cui, "A New Iris Segmentation Method for Recognition," in *Proceedings of the 17th IEEE International Conference on Pattern Recognition*, 2004, vol. 3, pp. 554–557.
- [67] S. Shah and A. Ross, "Iris Segmentation using Geodesic Active Contours," *IEEE Transactions on Information Forensics and Security*, vol. 4, no. 4, pp. 824–836, December 2009.
- [68] A. Abhyankar and S. Schuckers, "Active Shape Models for Effective Iris Segmentation," in *Proceedings of SPIE*, 2006, vol. 6202, pp. 166–175.
- [69] S. Pundlik, D. Woodard, and S. Birchfield, "Iris Segmentation in Non-Ideal Images using Graph Cuts," *Image and Vision Computing*, vol. 28, no. 12, pp. 1671–1681, 2010.

- [70] X. Hu, V. Pauca, and R. Plemmons, "Iterative Directional Ray based Iris Segmentation for Challenging Periocular Images," in *Biometric Recognition*, Zhenan Sun, Jianhuang Lai, Xilin Chen, and Tieniu Tan, Eds., vol. 7098 of *Lecture Notes in Computer Science*, pp. 91–99. Springer Berlin Heidelberg, 2011.
- [71] F. Alonso-Fernandez, P. Tome-Gonzalez, V. Ruiz-Albacete, and J. Ortega-Garcia, "Iris Recognition based on SIFT Features," in *International Conference on Biometrics, Identity and Security (BIDS)*, September 2009, pp. 1–8.
- [72] R. Jillela and A. Ross, "Methods for Iris Segmentation," in *Handbook of Iris Recognition*, K Bowyer and M. Burge, Eds. Springer, 2012.
- [73] R. Jillela, A. Ross, V. Boddeti, B. VijayaKumar, X. Hu, R. Plemmons, and P. Pauca, "An Evaluation of Iris Segmentation Algorithms in Challenging Periocular Images," in *Handbook of Iris Recognition*, K Bowyer and M. Burge, Eds. Springer, 2012.
- [74] J. Daugman, "The Importance of Being Random: Statistical Principles of Iris Recognition," *Pattern Recognition*, vol. 36, no. 2, pp. 279–291, 2003.
- [75] N. Kalka, N. Bartlow, and B. Cukic, "An Automated Method for Predicting Iris Segmentation Failures," in *IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems*. IEEE, 2009, pp. 1–8.
- [76] X. Liu, K. Bowyer, and P. Flynn, "Experiments with an Improved Iris Segmentation Algorithm," in *Fourth IEEE Workshop on Automatic Identification Advanced Technologies*. IEEE, 2005, pp. 118–123.
- [77] L. Masek and P. Kovesi, "MATLAB Source Code for a Biometric Identification System Based on Iris Patterns," 2003.
- [78] H. Proença, "Iris Recognition: On the Segmentation of Degraded Images Acquired in the Visible Wavelength," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 8, pp. 1502–1516, 2010.
- [79] T. Tan, X. Zhang, Z. Sun, and H. Zhang, "Noisy Iris Image Matching by Using Multiple Cues," *Pattern Recognition Letters*, 2011.
- [80] H. Proença and L. A. Alexandre, "The NICE I: Noisy Iris Challenge Evaluation," in *First IEEE International Conference on Biometrics: Theory, Applications, and Systems*. IEEE, 2007, pp. 1–4.
- [81] K. Bae, S. Noh, and J. Kim, "Iris Feature Extraction using Independent Component Analysis," in *Audio-and Video-Based Biometric Person Authentication*. Springer, 2003, pp. 1059–1060.
- [82] V. Dorairaj, N. Schmid, and G. Fahmy, "Performance Evaluation of Iris based Recognition System implementing PCA and ICA Encoding Techniques," in *Defense and Security*. International Society for Optics and Photonics, 2005, pp. 51–58.
- [83] Z. He, Z. Sun, T. Tan, X. Qiu, C. Zhong, and W. Dong, "Boosting Ordinal Features for Accurate and Fast Iris Recognition," in *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2008, pp. 1–8.

- [84] Z. He, T. Tan, Z. Sun, and X. Qiu, "Toward Accurate and Fast Iris Segmentation for Iris Biometrics," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 9, pp. 1670–1684, 2009.
- [85] X. Yuan and P. Shi, "Iris Feature Extraction using 2D Phase Congruency," in *Third International Conference on Information Technology and Applications*, July 2005, vol. 2, pp. 437–441.
- [86] F. Alonso-Fernandez, P. Tome-Gonzalez, V. Ruiz-Albacete, and J. Ortega-Garcia, "Iris Recognition based on SIFT Features," in *International Conference on Biometrics, Identity and Security (BIDS)*. IEEE, 2009, pp. 1–8.
- [87] J. Daugman, "Demodulation by Complex-Valued Wavelets for Stochastic Pattern Recognition," *International Journal of Wavelets, Multi-resolution and Information Processing*, vol. 1, pp. 1–17, 2003.
- [88] National Institute of Standards and Technology, "Iris Exchange (IREX)," 2007-2012.
- [89] S. Bhat and M. Savvides, "Evaluating Active Shape Models for Eye Shape Classification," in *IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2008, pp. 5228–5231.
- [90] U. Park, A. Ross, and A. Jain, "Periocular Biometrics in the Visible Spectrum: A Feasibility Study," in *IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems*. IEEE, 2009, pp. 1–6.
- [91] S. Bharadwaj, H. Bhatt, M. Vatsa, and R. Singh, "Periocular Biometrics: When Iris Recognition Fails," in *Biometrics: Theory, Applications and Systems*, September 2010.
- [92] V. P. Pauca, M. Forkin, X. Xu, R. Plemmons, and A. Ross, "Challenging Ocular Image Recognition," in *SPIE*, April 2011, vol. 8029, pp. 80291V–1–80291V–13.
- [93] J. Merkow, B. Jou, and M. Savvides, "An Exploration of Gender Identification using only the Periocular Region," in *Fourth IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS)*, September 2010, pp. 1–5.
- [94] K. Hollingsworth, S. Darnell, P. Miller, D. Woodard, K. Bowyer, and P. Flynn, "Human and Machine Performance on Periocular Biometrics under Near-Infrared Light and Visible Light," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, pp. 588–601, April 2012.
- [95] K. Hollingsworth, K. Bowyer, and P. Flynn, "Identifying Useful Features for Recognition in Near-Infrared Periocular Images," in *Fourth IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS)*, September 2010, pp. 1–8.
- [96] V. Boddeti, J. Smereka, and B. Vijaya Kumar, "A Comparative Evaluation of Iris and Ocular Recognition Methods on Challenging Ocular Images," in *International Joint Conference on Biometrics (IJCB)*, October 2011, pp. 1–8.

- [97] J. Daugman and C. Downing, "Effect of Severe Image Compression on Iris Recognition Performance," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 1, pp. 52–61, 2008.
- [98] A. Ross, K. Nandakumar, and A. Jain, *Handbook of Multibiometrics*, Springer Publishers, 2006.
- [99] G. Fahmy, "Super-Resolution Construction of Iris Images from a Visual Low Resolution Face Video," *National Radio Science Conference*, pp. 1–6, March 2007.
- [100] J. Huang, L. Ma, T. Tan, and Y. Wang, "Learning based Resolution Enhancement of Iris Images," *British Machine Vision Conference*, pp. 153–162, 2003.
- [101] K. Hollingsworth, T. Peters, K. Bowyer, and P. Flynn, "Iris Recognition using Signal-Level Fusion of Frames from Video," *IEEE Transactions on Information Forensics and Security*, vol. 4, no. 4, pp. 837–848, December 2009.
- [102] M. Turk and A. Pentland, "Eigenfaces for Recognition," *Journal of cognitive neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.
- [103] NIST, "Multi Biometric Grand Challenge," <http://face.nist.gov/mbgc/>, 2007.
- [104] NIST, "Iris Challenge Evaluation (ICE)," <http://iris.nist.gov/ice/>, 2006.
- [105] Neurotechnology, "VeriEye Iris Recognition Software," <http://www.neurotechnology.com/verieye.html/>, 2009.
- [106] K. Venkataramani and B. Vijaya Kumar, "Fingerprint Verification using Correlation Filters," in *Audio-and Video-Based Biometric Person Authentication*. Springer, 2003, pp. 886–894.
- [107] R. Gonzalez, R. Woods, and S. Eddins, *Digital Image Processing using MATLAB*, vol. 2, 2009.
- [108] P. Grother, G. Quinn, and P. Phillips, "Report on the Evaluation of 2D Still Image Face Recognition Algorithms," *National Institute of Standards and Technology Interagency Internal Report (NISTIR)*, vol. 7709, 2010.
- [109] J. Phillips, J. Beveridge, B. Draper, G. Givens, A. O'Toole, D. Bolme, J. Dunlop, Y. Lui, H. Sahibzada, and S. Weimer, "An Introduction to the Good, the Bad, & the Ugly Face Recognition Challenge Problem," in *IEEE International Conference on Automatic Face & Gesture Recognition and Workshops*. IEEE, 2011, pp. 346–353.
- [110] A. Jain, B. Klare, and U. Park, "Face Recognition: Some Challenges in Forensics," in *IEEE International Conference on Automatic Face & Gesture Recognition and Workshops*. IEEE, 2011, pp. 726–733.
- [111] American Society of Plastic Surgeons, "The History of Plastic Surgery," <http://www.plasticsurgery.org>, 2012.
- [112] American Society of Plastic Surgeons, "Statistics of Cosmetic Procedures," <http://www.plasticsurgery.org/news-and-resources/2012-plastic-surgery-statistics.html>, 2013.

- [113] American Society of Plastic Surgeons, “Statistics of Facial Plastic Surgeries,” <http://www.plasticsurgery.org/Documents/news-resources/statistics/2012-Plastic-Surgery-Statistics/Top-5-Cosmetic-Surgery-Procedures-2012.pdf>, 2013.
- [114] R. Singh, M. Vatsa, H. Bhatt, S. Bharadwaj, A. Noore, and S. Nooreyezdian, “Plastic Surgery: A New Dimension to Face Recognition,” *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 3, pp. 441–448, 2010.
- [115] W. Jones, “Plastic Surgery 1, Face Recognition 0,” *IEEE Spectrum*, vol. 46, no. 9, pp. 17, September 2009.
- [116] H. Bhatt, S. Bharadwaj, R. Singh, M. Vatsa, and A. Noore, “Evolutionary Granular Approach for Recognizing Faces Altered due to Plastic Surgery,” *IEEE International Conference on Automatic Face Gesture Recognition and Workshops*, pp. 720–725, March 2011.
- [117] G. Aggarwal, S. Biswas, P. Flynn, and K. Bowyer, “A Sparse Representation Approach to Face Matching Across Plastic Surgery,” *IEEE Workshop on Applications of Computer Vision*, January 2012.
- [118] M. De Marsico, M. Nappi, D. Riccio, and H. Wechsler, “Robust Face Recognition after Plastic Surgery using Local Region Analysis,” *Proceedings of International Journal of Computer Vision*, vol. 6754, pp. 191–200, 2011.
- [119] P. Miller, A. Rawls, S. Pundlik, and D. Woodard, “Personal Identification Using Periocular Skin Texture,” in *Proceedings of the 2010 ACM Symposium on Applied Computing*, 2010, pp. 1496–1500.
- [120] Neurotechnology, “Verilook 3.2,” <http://www.neurotechnology.com>, 2012.
- [121] PittPatt, “Pittsburgh Pattern Recognition,” now acquired by Google, 2012.
- [122] P. Viola and M. Jones, “Rapid Object Detection using a Boosted Cascade of Simple Features,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, 2001, vol. 1, pp. 511–518.
- [123] D. Lowe, “Distinctive Image Features from Scale-Invariant Key Points,” *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [124] T. Ojala, M. Pietikainen, and T. Maenpaa, “Multi-resolution Gray-scale and Rotation Invariant Texture Classification with Local Binary Patterns,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.
- [125] A. Vedaldi and B. Fulkerson, “VLFeat: An Open and Portable Library of Computer Vision Algorithms,” <http://www.vlfeat.org/>, 2008.
- [126] A. Ross, K. Nandakumar, and A. Jain, *Handbook of Multibiometrics*, vol. 6, Springer, 2006.

- [127] R. Jillela, A. Ross, and P. Flynn, “Information Fusion in Low-resolution Iris Videos using Principal Components Transform,” in *IEEE Workshop on Applications of Computer Vision*. IEEE, 2011, pp. 262–269.
- [128] R. Connaughton, A. Sgroi, K. Bowyer, and P. Flynn, “A cross-sensor evaluation of three commercial iris cameras for iris biometrics,” in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2011 IEEE Computer Society Conference on*. IEEE, 2011, pp. 90–97.
- [129] L. Xiao, Z. Sun, and T. Tan, “Fusion of Iris and Periocular Biometrics for Cross-Sensor Identification,” in *Biometric Recognition*, pp. 202–209. Springer, 2012.
- [130] C. Tan and A. Kumar, “Towards Online Iris and Periocular Recognition under Relaxed Imaging Constraints,” in *IEEE Transactions on Image Processing*, 2013.
- [131] F. Juefei-Xu, K. Luu, M. Savvides, T. Bui, and C. Suen, “Investigating Age Invariant Face Recognition based on Periocular Biometrics,” in *International Joint Conference on Biometrics*. IEEE, 2011, pp. 1–7.
- [132] V. Boddeti, J. Smereka, and B. Vijaya Kumar, “A Comparative Evaluation of Iris and Ocular Recognition Methods on Challenging Ocular Images,” in *International Joint Conference on Biometrics*, October 2011.
- [133] Y. Dong and D. Woodard, “Eyebrow Shape-based Features for Biometric Recognition and Gender Classification: A Feasibility Study,” in *International Joint Conference on Biometrics*. IEEE, 2011, pp. 1–8.
- [134] R. Jillela and A. Ross, “Mitigating Effects of Plastic Surgery: Fusing Face and Ocular Biometrics,” in *International Conference on Biometrics: Theory, Applications and Systems*. IEEE, 2012, pp. 402–411.
- [135] C. Padole and H. Proenca, “Periocular Recognition: Analysis of Performance Degradation Factors,” in *IAPR International Conference on Biometrics*. IEEE, 2012, pp. 439–445.
- [136] B. Oh, K. Oh, and K. Toh, “On Projection-based Methods for Periocular Identity Verification,” in *IEEE Conference on Industrial Electronics and Applications*, 2012, pp. 871–876.
- [137] K. Hollingsworth, S. Clark, J. Thompson, P. Flynn, and K. Bowyer, “Eyebrow Segmentation using Active Shape Models,” in *SPIE Defense, Security, and Sensing*. International Society for Optics and Photonics, 2013, pp. 871208–871208.
- [138] N. Dalal and B. Triggs, “Histograms of Oriented Gradients for Human Detection,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, 2005, vol. 1, pp. 886–893.
- [139] M. Pietikainen, “Image Analysis with Local Binary Patterns,” *Image Analysis, Lecture Notes in Computer Science*, vol. 3540, pp. 115–118, 2005.

- [140] G. Tzimiropoulos, V. Argyriou, S. Zafeiriou, and T. Stathaki, “Robust FFT-based Scale Invariant Image Registration with Image Gradients,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 10, pp. 1899–1906, 2010.
- [141] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, “Robust Face Recognition via Sparse Representation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210–227, 2009.
- [142] J. Pillai, V. Patel, R. Chellappa, and N. Ratha, “Secure and Robust Iris Recognition using Random Projections and Sparse Representations,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 9, pp. 1877–1893, 2011.
- [143] A. Kumar and T. Chan, “Robust Ear Identification Using Sparse Representation of Local Texture Descriptors,” *Pattern Recognition*, vol. 46, no. 1, pp. 73–85, January 2013.
- [144] M. Elad and M. Aharon, “Image denoising via sparse and redundant representations over learned dictionaries,” *Image Processing, IEEE Transactions on*, vol. 15, no. 12, pp. 3736–3745, 2006.
- [145] Julien Mairal, Guillermo Sapiro, and Michael Elad, “Learning multiscale sparse representations for image and video restoration,” Tech. Rep., DTIC Document, 2007.
- [146] J. Yang, J. Wright, T. Huang, and Y. Ma, “Image Super-resolution as Sparse Representation of Raw Image Patches,” in *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2008, pp. 1–8.
- [147] H. Guo, R. Wang, J. Choi, and L. Davis, “Face Verification using Sparse Representations,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. IEEE, 2012, pp. 37–44.
- [148] I. Tosic and P. Frossard, “Dictionary Learning,” *IEEE Signal Processing Magazine*, vol. 28, no. 2, pp. 27–38, 2011.
- [149] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, “Online Dictionary Learning for Sparse Coding,” in *Proceedings of the 26th Annual International Conference on Machine Learning*. ACM, 2009, pp. 689–696.
- [150] H. Lee, A. Battle, R. Raina, and A. Ng, “Efficient Sparse Coding Algorithms,” in *Advances in neural information processing systems*, 2006, pp. 801–808.
- [151] J. Yang, J. Wright, T. Huang, and Y. Ma, “Image Super-resolution via Sparse Representation,” *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 2861–2873, 2010.
- [152] S. Shekhar, V. Patel, N. Nasrabadi, and R. Chellappa, “Joint Sparsity based Robust Multimodal Biometrics Recognition,” in *European Conference on Computer Vision (ECCV) Workshop*. Springer, 2012, pp. 365–374.
- [153] R. Wildes, J. Asmuth, G. Green, S. Hsu, R. Kolczynski, J. Matey, and S. McBride, “A System for Automated Iris Recognition,” in *Proceedings of the Second IEEE Workshop on Applications of Computer Vision*. IEEE, 1994, pp. 121–128.

- [154] J. Illingworth and J. Kittler, “A Survey of the Hough Transform,” *Computer Vision, Graphics, and Image Processing*, vol. 44, no. 1, pp. 87–116, 1988.
- [155] H. Proença, “Iris Recognition: A Method to Segment Visible Wavelength Iris Images Acquired On-the-move and At-a-distance,” in *Advances in Visual Computing*, pp. 731–742. Springer, 2008.